SMOOTHED AGGREGRATION MULTIGRID FOR THE DISCONTINUOUS GALERKIN METHOD

F. PRILL[†‡§], M. LUKÁČOVÁ-MEDVIĎOVÁ[‡], AND R. HARTMANN[†]

Abstract. The aim of this paper is to investigate theoretically as well as experimentally an algebraic multilevel algorithm for the solution of the linear systems arising from the discontinuous Galerkin method. The smoothed aggregation multigrid, introduced by Vaněk for the conforming finite element method, is applied to low-order discretizations of convection-diffusion equations. For the elliptic model problem the algorithm is shown to be optimal. Adjustments for the case of non-vanishing advection, such as directionally implicit smoothing and a suitable splitting of the operator are discussed. Several numerical experiments are presented, including a Newton type linearization of the compressible Navier-Stokes equations.

Key words. Discontinuous Galerkin Method, Multigrid Method, Advection-Diffusion Equation, Navier-Stokes Equations.

AMS subject classifications. 35J05, 35J25, 35Q30, 65F10, 65M22, 65M55, 65N30, 65N35.

1. Introduction. The discontinuous Galerkin method (dGFEM) has become popular for the discretization of elliptic and hyperbolic partial differential equations for about a decade. To the desirable properties of this finite element method belong the flexibility in handling unstructured triangulations of complex geometries, mesh adaptation and freedom in the choice of the polynomial basis. In addition, the dGFEM can be viewed as a generalization of the finite volume method and recent studies have shown its suitability for the simulation of compressible flows [4, 19]. Moreover, the discretization possesses the variational background of the finite element method and is amenable to error analysis [2].

A major drawback is the fact that the linear algebraic systems arising from the dGFEM discretization consist of a relatively large number of unknowns and exhibit the customary numerical difficulties such as an increasing stiffness on large, anisotropic meshes. This motivates the development of efficient and scalable solvers, mostly belonging to the field of domain decomposition methods [1, 12, 27] or related defect-correction approaches, for example the so-called p-multilevel algorithms [13]. These approaches reduce the systems to low-order subproblems that are usually discretized with piecewise constant or linear polynomials. It is the point of this paper to propose an optimal and scalable iterative scheme for these remaining algebraic systems arising from dGFEM discretizations of low order. We recall that the convergence rate of a (quasi-)optimal algorithm increases at most logarithmically with the size of a problem. The method is scalable, if also the computational work for a problem containing n unknowns grows like $\mathcal{O}(n \log n)$.

Multigrid methods are known to be scalable for elliptic problems on regular grids. On structured meshes the grid hierarchy defined by successive global refinement steps can conveniently be employed for a multilevel method. The situation is different for unstructured grids and/or irregular domains. Here a mesh hierarchy is not provided *a priori* and an alternative coarse level construction must be developed.

One possibility is to utilize a (not necessarily nested) family of triangulations of

[†]Institute of Aerodynamics and Flow Technology, German Aerospace Center (DLR), Braunschweig, Germany.

[‡]Institute of Numerical Simulation, Hamburg University of Technology, Hamburg, Germany.

[§]Corresponding author, e-mail: florian.prill@dlr.de.

increasing grid size that has been generated automatically from the given geometry. However, especially in three dimensions automatic mesh generation is not a simple task. The second possibility is to construct the variational problems on grids that are defined purely algebraically. These approaches are termed *algebraic multigrid methods* (AMG) though besides the classical AMG method introduced in the early 1980s there exists a variety of other approaches exploiting the matrix data, such as AMGe [23], the AMLI method or (smoothed) aggregation multigrid [35, 36], the latter being the focus of this work.

The basic idea of smoothed aggregation multigrid is fairly intuitive and related to the geometric definition of coarser meshes. It consists in the construction of a non-overlapping partition of the domain of interest, where each of the subsets constitutes a node on the next coarser level. The characteristic functions associated with this partition, however, are not suitable for the definition of a coarse basis. Therefore, their properties are improved by a smoothing procedure. Aggregation multigrid is closely related to the agglomeration multigrid algorithm which is in widespread use for the finite volume method [29]. This makes the approach especially attractive for the discontinuous Galerkin methods applied to convection-diffusion systems, where the hyperbolic terms are discretized by numerical fluxes as well.

For multigrid by smoothed aggregation restricted to elliptic model problems, there exists a convergence theory proving the optimality of the algorithm in terms of the mesh size with bounds on the condition number that are only polynomial in the number of levels. It is based on the variational multigrid theory of Bramble [5, 6] and does not require regularity assumptions, in contrast to classical convergence proofs for geometric multigrid [17]. The theory is closely related to the abstract convergence proofs for additive or multiplicative Schwarz methods [33].

In general, literature of the application of multilevel methods in the context of dGFEM is still not comprehensive. While being widely used and well analyzed for the conforming finite element method, see for example [8] and the references cited therein, multigrid for the dGFEM has been described so far only by few researchers, see [7, 9, 14] for the geometric variant, and [25] for a formulation of the AMLI method. Additionally, a less rigorous study using local Fourier analysis has been performed and published by Hemker and co-workers [21], apart from some results published for engineering applications.

Another research focus has been the analysis of multilevel Schwarz methods. Similar to multigrid development most coarse problems have been either based on meshes that are provided *a priori* or they are constructed as low-order discretizations while keeping the triangulation unchanged [12, 27]. An exception is the report by Lasser et al. [26], where the results for a two-level method utilizing an algebraic coarse problem have been presented.

In this paper we describe and analyze theoretically a smoothed aggregation multigrid for the dGFEM. This algorithm can be applied to elliptic problems on unstructured meshes. Additionally, we enrich it by components from geometric and finite volume multigrid, such as an aggregation heuristic based on strong couplings and a line-implicit smoothing iteration. We apply the regularity-free convergence theory of Vaněk and Guillard to the dGFEM discretization and prove optimality of the solver in the elliptic case. In particular, we use a Petrov-Galerkin variant that has been proposed in [15, 16]. This method is well suited for combination with standard agglomeration multigrid and less expensive in terms of memory requirements. The scalability is demonstrated by numerous experiments on structured and unstruc-

tured quadrilateral meshes. A comparison with standard geometric multigrid will be performed as well where this is possible.

In addition to the model case of isotropic diffusion, we also apply the algorithm to non-symmetric problems using a suitable splitting of the problem operator. We consider linear convection-diffusion and test the feasibility of the method as a preconditioner for the Newton-type implicit linearizations of nonlinear convection-diffusion. A typical benchmark are the two-dimensional compressible Navier-Stokes equations modelling the flow around a NACA 0012 airfoil.

The paper is structured as follows: In Section 2 we briefly state the discontinuous Galerkin discretization of a scalar convection-diffusion model problem, along with some notation and the definition of the finite element spaces. Then, in Section 3, we formulate the smoothed aggregation multigrid of Vaněk and co-workers and discuss some adjustments to the dGFEM model problem such as line-implicit smoothing and a separate treatment of the convective and diffusive terms. Section 4 is devoted to the convergence results of multigrid theory applied to the dGFEM. We present some numerical experiments in Section 5 demonstrating the feasibility of the approach for purely diffusive and convection-diffusion problems. Finally, we draw conclusions in Section 6.

2. Model problem and discretization. In the following we state the model problem and its discretization. We consider stabilization by interior penalties and by the method of Bassi and Rebay [4].

2.1. Model problem. Let us consider the linear convection-diffusion equation on a bounded open polyhedral domain $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, with its boundary Γ being the union of its $(d-1)$-dimensional faces:

$$\mathcal{L}\, u \equiv -\nabla \cdot (\underline{a} \nabla u) + \beta \cdot \nabla u = f \quad \text{in } \Omega. \tag{2.1}$$

We assume that $f \in L^2(\Omega)$ and $\underline{a} \in \mathbb{R}^{d\times d}$ is a symmetric positive-definite matrix, and $\beta = \{\beta_i\}_{i=1}^d$ is a vector field in $W^{1,\infty}(\Omega)$ such that $\nabla \cdot \beta \leq 0$ in Ω. The existence of a unique solution for problem (2.1) equipped with homogeneous boundary conditions is well-known, e. g., [32].

In the general case, the boundary Γ is partitioned into sets

$$\begin{aligned} \Gamma_0 &:= \{\mathbf{x} \in \Gamma \ : \ \mathbf{n}(\mathbf{x})^T \underline{a}(\mathbf{x}) \mathbf{n}(\mathbf{x}) > 0\}, \\ \Gamma_- &:= \{\mathbf{x} \in \Gamma \setminus \Gamma_0 \ : \ \beta(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) < 0\}, \\ \Gamma_+ &:= \{\mathbf{x} \in \Gamma \setminus \Gamma_0 \ : \ \beta(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \geq 0\}, \end{aligned} \tag{2.2}$$

where $\mathbf{n}(\mathbf{x})$, $\mathbf{x} \in \Gamma$, denotes the unit outward normal vector and Γ_-, Γ_+ the inflow and outflow part of the boundary, respectively. The boundary Γ_0 is further divided into sets $\Gamma_0 = \Gamma_\mathrm{D} \cup \Gamma_\mathrm{N}$, such that the Dirichlet and Neumann boundary conditions can be imposed as

$$u = g_\mathrm{D} \text{ on } \Gamma_\mathrm{D} \cup \Gamma_-, \quad \mathbf{n} \cdot (\underline{a} \nabla u) = g_\mathrm{N} \text{ on } \Gamma_\mathrm{N},$$

assuming $\beta \cdot \mathbf{n} \geq 0$ on Γ_N.

2.2. Meshes, trace operators and finite element spaces. Let us assume that Ω can be subdivided into shape-regular meshes $\mathcal{T}_h$, consisting of convex open

subsets (elements) $\kappa_j \neq \emptyset$, $j = 1, \ldots, n_t$, of characteristic size $h > 0$:

$$\bar{\Omega} = \bigcup_{\kappa_j \in \mathcal{T}_h} \bar{\kappa}_j, \quad \kappa_i \cap \kappa_j = \emptyset, \quad 1 \leq i, j \leq n_t, i \neq j, \quad n_t := \operatorname{card} \mathcal{T}_h.$$

Throughout this paper we confine ourselves to partitionings into quadrilateral elements. We define the set $\mathcal{E}$ of all interior and boundary faces, i. e. the smallest $(d-1)$-dimensional intersection between neighboring elements of the partition and between elements and the boundary Γ. Furthermore, we define the set of interior faces $\mathcal{E}_{\text{int}} = \{e \in \mathcal{E} : e \subset \Omega\}$ and its union $\Gamma_{\text{int}} := \{\mathbf{x} \in \Omega : \exists e \in \mathcal{E}_{\text{int}} \text{ with } \mathbf{x} \in e\}$. The dGFEM admits triangulations containing irregular vertices, also called hanging nodes, in a natural manner.

We define parametric finite elements based on the tesselation $\mathcal{T}_h$ in the usual way, see for example Quarteroni [32] for details, however without continuity constraints on the inter-element boundaries. The global discrete function space is given by

$$V_h^p := \{v \in L^2(\Omega) : v\big|_\kappa \circ \sigma_\kappa \in \mathcal{Q}_p(\hat{\kappa}) \ \forall \kappa \in \mathcal{T}_h\},$$

where $\sigma_\kappa : \hat{\kappa} \to \kappa$ denotes a sufficiently smooth mapping from a reference element $\hat{\kappa} := [-1, 1]^d$ to the triangulation and $\mathcal{Q}_p(\hat{\kappa})$ is the set of polynomials of degree less than or equal to p in each variable. We limit the focus of this paper to approximations with uniform polynomial degree in each element. As we are dealing with polynomials of low degree (piecewise (bi-)linear or constant) the specific choice of a basis for $\mathcal{Q}_p(\hat{\kappa})$, $n_\kappa := \dim(\mathcal{Q}_p(\hat{\kappa})) = (p+1)^d$, is less important now.

By $H^s(\mathcal{T}_h)$, $s \in \mathbb{R}^+$, we denote the broken Sobolev space, i. e. the space of functions on $\mathcal{T}_h$ whose restriction to an element $\kappa \in \mathcal{T}_h$ belong to $H^s(\kappa)$. The associated broken norms and seminorms are defined as $\|u\|^2_{s,\mathcal{T}_h} = \sum_{\kappa \in \mathcal{T}_h} \|u\|^2_{s,\kappa}$, $|u|^2_{s,\mathcal{T}_h} = \sum_{\kappa \in \mathcal{T}_h} |u|^2_{s,\kappa}$.

For a function $v \in H^1(\mathcal{T}_h)$ we define v^+_κ, $\kappa \in \mathcal{T}_h$, to be the inner trace of v on $\partial\kappa$. The traces of functions in $H^1(\mathcal{T}_h)$ belong to the vector space

$$T(\Gamma \cup \Gamma_{\text{int}}) := \prod_{\kappa \in \mathcal{T}_h} L^2(\partial\kappa)$$

and are double-valued for $\mathbf{x} \in \Gamma_{\text{int}}$, while on the boundary, $\mathbf{x} \in \Gamma$, the value $v(\mathbf{x})$ is unambiguous. For $\kappa \in \mathcal{T}_h$ with $\partial\kappa \setminus \Gamma \neq \emptyset$ there exists a neighboring element $\kappa' \in \mathcal{T}_h$ that shares a common edge $e = \bar{\kappa} \cap \bar{\kappa}' \in \mathcal{E}_{\text{int}}$ with κ. The outer trace v^-_κ of v on e is defined as the inner trace v^+_κ relative to the element κ'.

Furthermore, let us define the following jump and average operators. For $v \in H^1(\mathcal{T}_h)$ the jump of v on the edge $e = \bar{\kappa} \cap \bar{\kappa}' \in \mathcal{E}_{\text{int}}$ is given by

$$[\![v]\!]_e : T(\Gamma \cup \Gamma_{\text{int}}) \to L^2(\Gamma \cup \Gamma_{\text{int}}), \quad [\![v]\!]_e = v^+_\kappa \mathbf{n}^+_\kappa + v^-_\kappa \mathbf{n}^-_\kappa,$$

where $\mathbf{n}^+_\kappa$ and $\mathbf{n}^-_\kappa$ denote the unit outward normal vectors to κ and κ', respectively. Additionally, the mean value of v on the edge e is defined as

$$\{v\}_e : T(\Gamma \cup \Gamma_{\text{int}}) \to L^2(\Gamma \cup \Gamma_{\text{int}}), \quad \{v\}_e = \frac{1}{2}\left(v^+_\kappa + v^-_\kappa\right).$$

For element boundaries $e \in \mathcal{E} \setminus \mathcal{E}_{\text{int}}$, that are part of the global domain boundary Γ, the boundary values are defined unambiguously. We set

$$[\![v]\!]_e = v^+ \mathbf{n}, \quad \{v\}_e = v^+ \text{ on } e.$$

Finally, we introduce a jump notation for the inflow part $\partial_- \kappa$ of an element boundary that is defined similar to (2.2) as $\partial_- \kappa = \{\mathbf{x} \in \partial\kappa \ : \ \beta(\mathbf{x}) \cdot \mathbf{n} < 0\}$. Then the jump of u across $\partial_- \kappa \setminus \Gamma$ is defined by $[v]_\kappa := v_\kappa^+ - v_\kappa^-$. For the sake of simplicity the subscripts κ and e are omitted in the following.

2.3. The discontinuous Galerkin discretization. Arnold et al. [2] presented a unified formulation for the discontinuous Galerkin discretization of second-order operators which allows a simultaneous treatment of both the interior penalty stabilization and the approach of Bassi and Rebay (*BR2 method*).

The discrete problem is given in the following way: Find $u_h \in V_h^p$ such that

$$B_1(u_h, v_h) + B_2(u_h, v_h) = \ell(v_h) \quad \forall v_h \in V_h^p, \tag{2.3}$$

with bilinear forms B_1, B_2 for the diffusion and convection part of problem (2.1), respectively. The latter uses a standard upwind term and is given by

$$\begin{aligned} B_2(u,v) := & \sum_{\kappa \in \mathcal{T}_h} \int_\kappa (\beta \cdot \nabla_h u)\, v \, d\mathbf{x} - \sum_{\kappa \in \mathcal{T}_h} \int_{\partial_- \kappa \setminus \Gamma} (\beta \cdot \mathbf{n}) [u] v^+ \, ds \\ & - \sum_{\kappa \in \mathcal{T}_h} \int_{\partial_- \kappa \cap \Gamma} (\beta \cdot \mathbf{n}) u^+ v^+ \, ds, \\ \ell(v) := & - \sum_{\kappa \in \mathcal{T}_h} \int_{\partial_- \kappa \cap (\Gamma_- \cup \Gamma_{\mathrm{D}})} (\beta \cdot \mathbf{n}) g_{\mathrm{D}} v^+ \, ds + \sum_{\kappa \in \mathcal{T}_h} \int_\kappa f v \, d\mathbf{x}, \end{aligned}$$

where we used the notation ∇_h for the broken gradient $(\nabla_h u)\big|_\kappa = \nabla \left(u\big|_\kappa\right)$, $\kappa \in \mathcal{T}_h$.

The primal formulation of the elliptic operator is the following:

$$\begin{aligned} B_1(u,v) := & \int_\Omega \underline{a} \nabla_h u \cdot \nabla_h v \, d\mathbf{x} + \int_{\Gamma \cup \Gamma_{\mathrm{int}}} (\llbracket \hat{u} - u \rrbracket \cdot \{\underline{a} \nabla_h v\} - \{\hat{\mathbf{q}}\} \cdot \llbracket v \rrbracket) \ ds \\ & + \int_{\Gamma_{\mathrm{int}}} (\{\hat{u} - u\} \cdot \llbracket \underline{a} \nabla_h v \rrbracket - \llbracket \hat{\mathbf{q}} \rrbracket \cdot \{v\}) \ ds, \end{aligned}$$

where we used vector and scalar numerical fluxes on the boundary of κ,

$$\begin{aligned} & \hat{\mathbf{q}} : H^2(\mathcal{T}_h) \times \left[H^1(\mathcal{T}_h)\right]^d \to [T(\Gamma \cup \Gamma_{\mathrm{int}})]^d, && \hat{\mathbf{q}}(u, \nabla u, \mathbf{n}) \approx \underline{a} \nabla u, \\ & \hat{u} : H^1(\mathcal{T}_h) \to T(\Gamma \cup \Gamma_{\mathrm{int}}), && \hat{u}(u, \mathbf{n}) \approx u. \end{aligned}$$

Various choices of $\hat{\mathbf{q}}$, $\hat{u}$ give rise to a collection of consistent and stable dGFEM variants [2], among them the interior penalty method in its symmetric (SIPG) and non-symmetric (NIPG) form, where we have

$$\begin{aligned} & \hat{u}_{\mathrm{SIPG}}(u) := \{u\}, && \hat{\mathbf{q}}_{\mathrm{SIPG}}(u, \nabla u, \mathbf{n}) := \{\underline{a} \nabla u\} - \sigma \llbracket u \rrbracket, \quad \text{and} \\ & \hat{u}_{\mathrm{NIPG}}(u) := \{u\} + \mathbf{n} \cdot \llbracket u \rrbracket, && \hat{\mathbf{q}}_{\mathrm{NIPG}}(u, \nabla u, \mathbf{n}) := \hat{\mathbf{q}}_{\mathrm{SIPG}}(u, \nabla u, \mathbf{n}) \quad \text{on } \Gamma_{\mathrm{int}}, \end{aligned}$$

with appropriate modification at the boundary. The term $\sigma = \sigma(|\underline{a}|, h, p)$ denotes a stabilizing penalty parameter and is chosen as $\sigma\big|_e = \delta p^2 h^{-1} \| \sqrt{\underline{a}} \mathbf{n}_e \|_{L^\infty(e)}$ with a given factor $\delta > 0$.

The (modified) scheme of Bassi and Rebay [4] uses the numerical fluxes

$$\hat{u}_{\mathrm{BR2}}(u) := \hat{u}_{\mathrm{SIPG}}(u), \quad \hat{\mathbf{q}}_{\mathrm{BR2}}(u, \nabla u, \mathbf{n}) := \{\underline{a} \nabla u\} + \eta_e \{r_e(\llbracket u \rrbracket)\}, \tag{2.4}$$

where η_e, $e \in \mathcal{E}_{\text{int}}$, is again a stabilizing constant and $r_e : \left[L^1(e)\right]^d \to \left[H^1(\mathcal{T}_h)\right]^d$ denotes a local lifting operator which is defined in a weak manner by

$$\int_\Omega r_e(\varphi) \cdot \tau \, d\mathbf{x} = -\int_e \varphi \cdot \{\underline{a}^T \tau\} \, ds \quad \forall \tau \in \left[H^1(\mathcal{T}_h)\right]^d, \varphi \in \left[L^1(e)\right]^d.$$

Using a local polynomial basis for $\mathcal{Q}_p(\hat{\kappa})$ together with the discretization (2.3), the discrete operator of the boundary value problem (2.1) is transformed into a stiffness matrix $\underline{A} \in \mathbb{R}^{n \times n}$, $n := \dim V_h^p = n_t n_\kappa$, while the functional $\ell(\cdot)$ takes the form of a right hand side vector $\mathbf{b} \in \mathbb{R}^n$. Thus we are interested in solving a system of linear algebraic equations,

$$\underline{A}\mathbf{x} = \mathbf{b}.$$

Later, for the self-adjoint and coercive linear elliptic model problem, we will make use of the induced energy norm $\|\mathbf{u}\|_{\underline{A}} = (\mathbf{u}, \mathbf{u})_{\underline{A}}$, where $(\mathbf{u}, \mathbf{v})_{\underline{A}} = \mathbf{u}^T \underline{A}\mathbf{v}$, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

3. Description of the algorithm. In the following we state the classical multigrid algorithm and describe the smoothed aggregation approach for unstructured tesselations. Finally, we discuss the agglomeration heuristic and the smoothing iteration.

3.1. Basic algorithm. The smoothed aggregation multigrid can be viewed as a standard variational multigrid. We state the abstract algorithm [5] as a linear iteration $\mathbf{x} \leftarrow MG(\mathbf{x}, \mathbf{b})$. This is equivalent to a preconditioner formulation,

$$\mathbf{x} \leftarrow \mathbf{x} - \underline{B}_{MG} \left(\underline{A}\mathbf{x}_n - \mathbf{b}\right), \quad \underline{B}_{MG}\mathbf{g} := MG(\mathbf{0}, \mathbf{g}). \tag{3.1}$$

For the sake of simplicity, we focus on the V-cycle algorithm, although extensions as the W-cycle or full multigrid can be useful and are straightforward in implementation. Further we assume the model problem to be of purely elliptic type. The case of a non-vanishing advection term will be discussed in Section 3.5.

We assume a sequence of finite-dimensional coarse spaces and operators

$$M_1 \subset M_2 \subset \cdots \subset M_J \equiv \mathbb{R}^n, \quad \left\{\underline{A}_j\right\}_{j=1}^J, \ \underline{A}_J \equiv \underline{A},$$

where M_1 denotes the coarsest problem space. Further we have prolongation operators $I_{k-1}^k : M_{k-1} \to M_k$ as well as restrictions $I_k^{k-1} : M_k \to M_{k-1}$, and smoothing iterations of fixed-point type

$$\mathbf{x} \leftarrow (\underline{I} - \underline{R}_k \underline{A}_k)\mathbf{x} + \underline{R}_k \mathbf{b} \tag{3.2}$$

with preconditioners $\underline{R}_k : M_k \to M_k$, $k = 2, \ldots, J$. Inductively, we define $MG(\mathbf{x}, \mathbf{b}) \equiv MG_J(\mathbf{x}^J, \mathbf{b}^J)$ as follows:

Algorithm 1: Multigrid algorithm $\mathbf{x}^l \leftarrow MG_l(\mathbf{x}^l, \mathbf{b}^l)$

$MG_1(\mathbf{x}^1, \mathbf{b}^1) := \underline{A}^{-1}\mathbf{b}^1$ (*direct solution*)

for $l > 1$: Let $\mathbf{x}^l$, $\mathbf{b}^l \in M_l$ be given.

begin

- Perform m iterations of $\mathbf{x}^l \leftarrow \left(\underline{I} - \underline{R}_l^T \underline{A}_l\right)\mathbf{x}^l + \underline{R}_l^T \mathbf{b}^l$ (*pre-smoothing*)
- Set $\mathbf{b}^{l-1} = I_k^{k-1}\left(\mathbf{b}^l - \underline{A}_l \mathbf{x}^l\right)$.
- $\mathbf{x}^l \leftarrow \mathbf{x}^l + I_{k-1}^k \mathbf{x}^{l-1}$, where $\mathbf{x}^{l-1} := MG_{l-1}(\mathbf{0}, \mathbf{b}^{l-1})$ (*coarse-grid correction*)
- Perform m iterations of $\mathbf{x}^l \leftarrow (\underline{I} - \underline{R}_l \underline{A}_l)\mathbf{x}^l + \underline{R}_l \mathbf{b}^l$ (*post-smoothing*)

end

The specific choice of the subspaces $\underline{M}_k$, $k = 1, \ldots, J$, as well as the level operators $\underline{A}_k$ and the transfer operators will be described next. The construction of the smoothing iteration $\underline{R}_k$ is discussed in Section 3.4.

3.2. Smoothed aggregation coarse spaces. Now, we describe the smoothed aggregation multigrid [35, 36]. Our starting point is a non-overlapping partition of the domain into strongly connected subsets a_{i_m}, where i_m denotes the index for a degree of freedom, corresponding to a vector in M_m. Formally identifying the elements $\kappa \in \mathcal{T}_h$ with the aggregation $\mathcal{A}_J$ on the finest level we construct a set of level-m-agglomerates as

$$\mathcal{A}_m := \{a_{i_m} \subseteq \Omega \ : \ i_m \in \mathcal{I}_{\mathcal{A}_m} \subseteq \mathbb{N}^*\},$$

where each element cluster $a_{i_m} \subseteq \Omega$ is the union of level-$(m+1)$-agglomerates

$$a_{i_m} = \bigcup_{j_{m+1} \in \mathcal{I}_{i_m}} a_{j_{m+1}},$$

s.t. the index sets $\mathcal{I}_{i_m}$ partition the index sets of the level-$(m+1)$-agglomerates $\mathcal{I}_{\mathcal{A}_{m+1}}$

$$\mathcal{I}_{i_m} \cap \mathcal{I}_{j_m} = \emptyset \ \forall i_m \neq j_m, \qquad \bigcup_{i_m \in \mathcal{I}_{\mathcal{A}_m}} \mathcal{I}_{i_m} = \mathcal{I}_{\mathcal{A}_{m+1}}.$$

For each of the disjoint clusters a_{i_m} the degrees of freedom are aggregated and give rise to a common degree of freedom (*supernode*) on the coarser level $(m-1)$. This is realized recursively and thus $n_k := \dim M_k = \operatorname{card} \mathcal{I}_{\mathcal{A}_k}$, $k = 1, \ldots, J$. The structure of the spaces M_k is thus determined by the agglomeration strategy, with the exception of the level $J-1$, where we have some freedom in organizing the unknowns of M_J.

The following notation will be of later use. We define the composite set $\mathcal{I}^J_{l,k}$ of fine level indices and the corresponding aggregate $a^J_{l,k}$ by

$$\mathcal{I}^J_{l,k} := \left\{ i_J \in \mathcal{I}_{\mathcal{A}_J} \ : \ i_s \in \mathcal{I}_{i_{s-1}}, s = l+1, \ldots, J, i_l = k \right\}, \quad a^J_{l,k} := \bigcup_{i_J \in \mathcal{I}^J_{l,k}} a_{i_J}.$$

The characteristic functions $\chi_{a_{i_m}}$, corresponding to the aggregates a_i on level m, span a linear space which can be used as the ansatz space for a Galerkin procedure forming the coarse solution $\tilde{u}_m$ for the problem

$$B_1(\tilde{u}_m, \tilde{v}_m) = f(\tilde{v}_m) \quad \forall \tilde{v}_m \in \operatorname{span}_{i \in \mathcal{I}_{\mathcal{A}_m}} \chi_{a_i}.$$

In view of the matrix formulation this defines the so-called tentative prolongation and restriction

$$\tilde{I}^k_{k-1} \in \mathbb{R}^{n_k \times n_{k-1}}, \quad \left(\tilde{I}^k_{k-1}\right)_{ij} := \chi_{a_j}(\mathbf{x}_i), \quad \tilde{I}^{k-1}_k := \left(\tilde{I}^k_{k-1}\right)^T,$$

where $\mathbf{x}_i$ denotes the geometric location of the i-th unknown in M_k. In fact, the coarse degrees of freedom are simply constructed as the sums of the fine level unknowns inside the element cluster. In order to construct the interpolation operators we restrict ourselves to orthonormalized zero-one prolongations, though more general choices are possible [35].

It can be shown by an argument on the matrix graph that the maximum number of nonzero entries per line in the coarse stiffness matrices

$$\underline{\tilde{A}}_k = \left(\tilde{I}_k^J\right)^T \underline{A}\tilde{I}_k^J, \text{ with } \tilde{I}_k^J := \tilde{I}_{J-1}^J \cdots \tilde{I}_{k+1}^{k+2}\tilde{I}_k^{k+1}, \quad k = 1, \ldots, J-1, \tag{3.3}$$

is bounded under the condition that the partitioning algorithm keeps the number of direct neighbors of the agglomerates approximately constant [16]. This assures that the number of nonzeros decreases with the coarsening factor that is necessary for the scalability of the algorithm.

The tentative transfer operators, however, do not imply a convergent multilevel algorithm. This fact has been observed experimentally [28], and it is supported by the regularity-free convergence theory. Besides, it is a common rule of thumb in engineering applications that piecewise constant transfer operators are insufficient for multigrid applied to second order elliptic problems [20].

To provide a stable coarse basis we apply symmetric smoothing polynomials $s^k : \mathbb{R}^{n_k \times n_k} \to \mathbb{R}^{n_k \times n_k}$, $1 \le k \le J$, to the piecewise constant transfer operators. Following Guillard et al. [16] we employ different smoothers s_L^k, s_R^k for the restriction and the prolongation, respectively,

$$I_k^{k-1} := \left(s_L^k(\underline{A}_k)\tilde{I}_{k-1}^k\right)^T, \quad I_{k-1}^k := s_R^k(\underline{A}_k)\tilde{I}_{k-1}^k.$$

Then, defining the coarse level matrices as

$$\underline{A}_{k-1} = I_k^{k-1}\underline{A}_k I_{k-1}^k$$

is equivalent to the Galerkin procedure $\underline{A}_{k-1} = \left[B_1(\phi_i^{k-1}, \psi_j^{k-1})\right]_{i,j=1}^{n_{k-1}}$ with coarse basis functions ϕ_i^{k-1}, ψ_j^{k-1}, which are inductively defined as

$$\begin{bmatrix} \phi_1^{k-1} \\ \vdots \\ \phi_{n_{k-1}}^{k-1} \end{bmatrix} = \left(\tilde{I}_{k-1}^k\right)^T s_L(\underline{A}_k) \begin{bmatrix} \phi_1^k \\ \vdots \\ \phi_{n_k}^k \end{bmatrix}, \quad \begin{bmatrix} \psi_1^{k-1} \\ \vdots \\ \psi_{n_{k-1}}^{k-1} \end{bmatrix} = \left(\tilde{I}_{k-1}^k\right)^T s_R(\underline{A}_k) \begin{bmatrix} \psi_1^k \\ \vdots \\ \psi_{n_k}^k \end{bmatrix}.$$

The smoothing polynomials produce overlap between the supports of the coarse functions; this is especially noteworthy for the dGFEM, where the tentative coarse functions possess no overlap at all. However, the smoothed coarse functions remain piecewise discontinuous.

We choose between two variants of transfer operator smoothers.

1. The (symmetric) *Ritz-Galerkin* multigrid is given by

$$s_L^k(\underline{A}_k) = s_R^k(\underline{A}_k) := \underline{I} - \frac{4}{3\lambda_k}\underline{A}_k, \tag{3.4}$$

where $\lambda_k \ge \varrho(\underline{A}_k)$ is an approximation for the spectral radius.

2. The *Petrov-Galerkin* approach is defined by

$$s_L^k(\underline{A}_k) := \underline{I}, \quad s_R^k(\underline{A}_k) := \underline{I} - \frac{1}{\lambda_k}\underline{A}_k. \tag{3.5}$$

Other choices for optimal smoothing polynomials are given in [15]. The total degree of the smoothing polynomials is closely related to the diameter (coarsening ratio)

of the agglomerates a_i. While an increased degree improves the smoothing effect, the overlap increases, too. Therefore fill-in optimality of the level matrices demands either a smaller degree or larger agglomerates. Thus the smoothing polynomial must be carefully chosen to match the coarsening ratio.

In the presence of additional convective terms yet another argument takes place. Traditional finite-volume multigrid is based on a coarsening ratio of 2 to resolve hyperbolic features of the solution. Thus, the presence of hyperbolic terms demands the smallest possible degree for the coarsening ratio [22].

The tentative operators $\tilde{I}_{k-1}^k$ as well as the smoothed prolongations I_{k-1}^k and the level matrices $\underline{A}_k$ are computed in a setup phase. In addition, with an orthonormalization procedure (diagonal scaling), we achieve that $\left(\tilde{I}_{k-1}^k\right)^T \left(\tilde{I}_{k-1}^k\right) = \underline{I}$, which is necessary for convergence, see Theorem 4.5.

For completeness, we explicitly define the hierarchy of coarse spaces M_k, $k = 1, \ldots, J-1$, for the variational multigrid [35, 15]. The space M_k is given as

$$M_k := \text{Range}\left(I_k^J\right), \quad k = 1, \ldots, J-1,$$

where for the Ritz-Galerkin variant, we define I_k^J analogously to (3.3) as

$$I_k^J = s_R^J(\underline{A}_J)\tilde{I}_{J-1}^J \cdots s_R^{k+1}(\underline{A}_{k+1})\tilde{I}_k^{k+1}, \quad k = 1, \ldots, J-1.$$

For the Petrov-Galerkin multigrid, we set

$$I_k^J = \left(s_R^J(\underline{A}_J)\right)^{\frac{1}{2}} \tilde{I}_{J-1}^J \cdots \left(s_R^{k+1}(\underline{A}_{k+1})\right)^{\frac{1}{2}} \tilde{I}_k^{k+1}, \quad k = 1, \ldots, J-1.$$

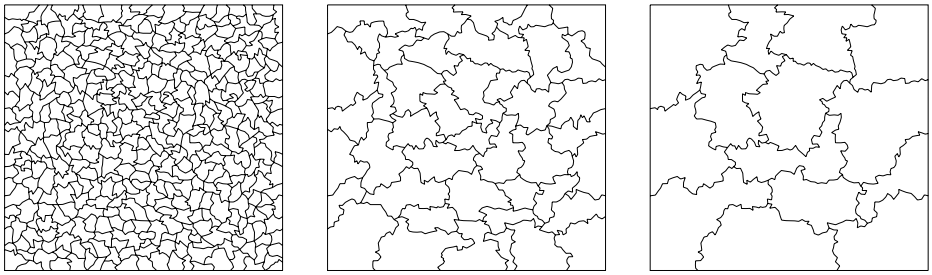
cf. Remark 1. M_k is equipped with the scalar product $(\cdot,\cdot)_k$, given by $(\mathbf{u},\mathbf{v})_k = (\mathbf{x},\mathbf{y})_{R^{n_k}}$, $\mathbf{u} = I_k^J\mathbf{x}$, $\mathbf{v} = I_k^J\mathbf{y}$, $\mathbf{x},\mathbf{y} \perp \mathcal{N}\left(I_k^J\right)$.

3.3. Agglomeration strategy. Partitionings of the domain into connected sets of neighboring elements are required by many hierarchical solution strategies, for example domain decomposition methods, agglomeration finite volume multigrid and algebraic multilevel approaches like AMGe. Thus, there exists a variety of coarsening algorithms in the literature [16, 22, 23, 24, 28, 30, 35].

Many approaches heuristically approximate a maximum independent set in the corresponding node graph and are of ideally linear complexity. In the current work we use a greedy clustering algorithm that operates on the face-face graph of the tesselation, i.e. the undirected graph whose vertices are faces of the triangulation with edges between faces sharing a common triangulation vertex. The algorithm was initially described by Jones and Vassilevski [23] in the context of the AMGe method. It can be extended to three-dimensional problems by operating on the side-side graph [24].

The basic idea behind the 2D algorithm is to identify each face of the triangulation with a weighting scalar that is incremented during the iterations whenever neighboring faces are processed. The algorithm then tries to build monotonically increasing sequences of face weights, while adding the adjacent elements to the new agglomerate.

The algorithm has a typical coarsening ratio of 2 in each spatial direction and it is able to rebuild the natural coarse hierarchy of $J = 1 - \log_2 h$ levels in globally refined structured meshes. On unstructured tesselations, however, the shape of the element clusters may deteriorate. Further, using this algorithm for the symmetric (Ritz-Galerkin) multigrid causes additional fill-in inside the level matrices. The total degree of the smoothing polynomial then requires a coarsening factor of 3 to keep the nearest neighbor stencil property.



(a) Level J-1 (b) Level J-3 (c) Level J-4

FIGURE 3.1. *Example of an agglomeration on an unstructured mesh with* 5316 *elements. The lines denote the boundaries of the composite aggregates* $a^J_{l,i}$ *of the agglomeration. See Figure 5.1(a) for a plot of the underlying tesselation.*

3.4. Smoothing iteration. In this section we discuss the multilevel smoothing iteration (3.2) in more detail. The operators $\underline{R}_k$, $k = 2, \dots, J$, were chosen to be of Jacobi or Gauß-Seidel type. In particular, a directionally implicit variant was implemented.

From an abstract point of view the so-called line-implicit relaxation schemes belong to the class of additive or multiplicative iterative schemes associated with a subspace decomposition $M_k = \sum_{i=1}^{l} M_k^i$. We assume projections $\underline{Q}_k^i : M_k \to M_k^i$ w.r.t. the inner product $(\cdot,\cdot)_k$. The additive (Jacobi) smoother is then defined by

$$\underline{R}_k = \gamma \sum_{i=1}^{l} \underline{A}_{k,i}^{-1} \underline{Q}_k^i, \quad k = 2, \dots, J,$$

with a damping factor $\gamma > 0$ and exact local problems $\underline{A}_{k,i} : M_k^i \to M_k^i$, $i = 1, \dots, l$, given by $(\underline{A}_{k,i}\mathbf{v}, \mathbf{w})_k = (\underline{A}_k\mathbf{v}, \mathbf{w})_k$ for all $\mathbf{w} \in M_k^i$. The multiplicative (Gauß-Seidel) variant is the result of the following procedure:

for $i = 1, \dots, l$ **do**
$\mathbf{v}_i = \mathbf{v}_{i-1} + \underline{A}_{k,i}^{-1} \underline{Q}_k^i (\mathbf{b} - \underline{A}_k \mathbf{v}_{i-1})$, where $\mathbf{v}_0 := \mathbf{0}$.
$\underline{R}_k \mathbf{b} = \mathbf{v}_l$.

For the case that the space of unknowns is decomposed into a direct sum, the above methods reduce to classical block Jacobi or Gauß-Seidel schemes applied to the stiffness matrix. Obviously the subspace decomposition determines the additive and multiplicative scheme.

The choice of the spaces is much less restricted for multilevel smoothers than for the related classical Schwarz methods. We set the following assumptions.

ASSUMPTION 1.

1. *We assume that the matrices* $\underline{R}_k$, $k = 2, \dots, J$, *satisfy*

$$\lambda_{\min}(\underline{R}_k) \geq \frac{1}{C_R^2 \varrho(\underline{A}_k)}, \quad k = 2, \dots, J,$$

with a constant $C_R > 0$ *independent of the level* k.

2. *We assume that there exists* $\theta \in [0, 2)$, *independent of the level* k, *s.t.*

$$(\underline{R}_k \underline{A}_k \mathbf{u}, \underline{R}_k \underline{A}_k \mathbf{u})_{\underline{A}_k} \leq \theta \, (\underline{R}_k \underline{A}_k \mathbf{u}, \mathbf{u})_{\underline{A}_k} \quad \forall \ \mathbf{u}_k \in \mathbb{R}^{n_k}.$$

For a symmetric positive definite smoother $\underline{R}_k$ *this assumption is satisfied with* $\theta = 1$*, when* $\lambda_{\min}\left(\underline{I} - \underline{R}_k\underline{A}_k\right) \geq 0$*, i.e. the spectrum lies within the set* $\sigma(\underline{I} - \underline{R}_k\underline{A}_k) \subseteq [0,1]$.

REMARK 1. *The Assumption 1 must be modified in the following way for the Petrov-Galerkin multigrid [16, 22]: The algorithm can be reformulated as a Ritz-Galerkin approach with smoothing matrix* $\tilde{s}^k(\underline{A}_k) := \left(s_R^k\right)^{\frac{1}{2}}(\underline{A}_k)$ *with the help of an additional pre-smoothing step*

$$\mathbf{x} \leftarrow \tilde{s}^k(\underline{A}_k)\mathbf{x} + (\underline{I} - \tilde{s}^k(\underline{A}_k))\underline{A}_k^+\mathbf{b}, \tag{3.6}$$

where $\underline{A}_k^+$ *denotes the pseudoinverse. Therefore, to apply the abstract convergence theory, Assumption 1 must hold for the modified smoother* $\underline{R}_k'$*,*

$$\underline{K}_k' = \tilde{s}^k(\underline{A}_k)(\underline{I} - \underline{R}_k\underline{A}_k), \quad \underline{R}_k' = (\underline{I} - \underline{K}_k')\underline{A}_k^+.$$

The additive and the multiplicative smoother both satisfy the Assumption 1 under weak conditions [6]. In particular, these are fulfilled when the triangulation is decomposed into *lines* of elements without loops or cycles. This is a popular smoothing scheme since the local solvers $\underline{A}_{k,i}^{-1}$ are (block) tridiagonal linear systems that can be efficiently factorized into an LU decomposition and solved via forward and backward substitution.

Whereas from the theory the particular way of line decomposition does not need to fulfil further properties, from a practical point of view it creates a difficult task. It is observed experimentally, and has been supported by local Fourier analysis [30], that lines should follow strong couplings in the stiffness matrix, when the iterative scheme is used as a multigrid smoother. In the case of pure convection the smoother then degrades into an exact solver. This is a common strategy for the construction of a robust multigrid method, that is ideally insensitive with respect to the strength $\|\beta\|_\infty$ of the advection term.

We make use of a coupling matrix $\underline{C} \in \mathbb{R}^{n_t \times n_t}$, $0 \leq c_{ij} = c_{ji} \leq 1$ for $1 \leq i, j \leq n_t$, which measures interdependencies between the elements of the triangulation. Elements $\kappa_i, \kappa_j \in \mathcal{T}_h$ with $c_{ij} \approx 1$ are termed *strongly coupled*. The matrix $\underline{C}$ is defined as the stiffness matrix $\underline{A}_{\tilde{\mathcal{L}}} = \left\{a_{ij}^{\tilde{\mathcal{L}}}\right\}_{i,j=1}^{n_t}$ for a scalar model problem (2.1) with operator $\tilde{\mathcal{L}}$, normalized by

$$\underline{C} = c_{ij}, \quad c_{ij} := \frac{\tilde{c}_{ij}}{\max_{1\leq i\leq n_t}\tilde{c}_{ij}}, \quad 1 \leq i, j \leq n_t,$$
$$\text{with } \tilde{c}_{ij} = \max\left(|a_{ij}^{\tilde{\mathcal{L}}}|, |a_{ji}^{\tilde{\mathcal{L}}}|\right) \approx \max\left(\left|\frac{\partial\tilde{\mathcal{L}}_i}{\partial\phi_j}\right|, \left|\frac{\partial\tilde{\mathcal{L}}_j}{\partial\phi_i}\right|\right).$$

The BR2 method is more appropriate for the discretization of $\tilde{\mathcal{L}}$ than the interior penalty variant. While the piecewise constant SIPG approximation of the diffusive operator reduces to a simple boundary integral with constant factor $\sigma = \sigma(|\underline{a}|, h, p)$, the stability term of the BR2 scheme still captures the directions of anisotropic diffusion through the formulation of the lifting operator r_e. When dealing with a linear primal problem (2.1) the same coefficients $\underline{a}$, β can be chosen for $\tilde{\mathcal{L}}$. The task is less trivial for nonlinear convection-diffusion. There, one can extract the vector β from the nonlinear state. However, the diffusion tensor $\underline{a}$ must be chosen *a priori*.

With an appropriate coupling matrix at hand we use a greedy line construction heuristic starting with seed elements to extract the subspaces M_k^i, $i = 1, \ldots, l$, from

the triangulation, where l denotes the number of lines. The algorithm is similar to the one described by Okusanya [13, 30]. A typical result is shown in Figure 5.1(b). The same strategy can be applied to the triangulation as well as to the agglomerated levels, where we define

$$\underline{C}^{k-1} = \tilde{I}_k^{k-1} \underline{C}^k \tilde{I}_{k-1}^k, \quad k = 2, \dots, J, \quad \underline{C}^J := \underline{C}.$$

3.5. Treatment of convective terms. The smoothed aggregation technique has been described and analyzed by Vaněk et al. for elliptic problems. The discretization of the convection-diffusion problem (2.1), however, includes also a hyperbolic term whose upwind character deteriorates under the smoothed interpolation. Following Guillard [16] we therefore split the discrete operator

$$\underline{A}\mathbf{x} = (\underline{A}^v + \underline{A}^c)\,\mathbf{x}$$

into its diffusive and convective components $\underline{A}^v$ and $\underline{A}^c$, respectively. Then, the coarse level matrices for the convective term are constructed using the simple, non-smoothed transfer operators $\tilde{I}_{k-1}^k$, $\tilde{I}_k^{k-1}$. It can be shown that aggregation with piecewise constant interpolation is a sufficient approximation of the rediscretized operator on the agglomerates.

4. Convergence bounds. The convergence estimate for the smoothed aggregation multigrid requires some theoretical tools that are presented in the Section 4.1. Throughout this section we consider an elliptic model problem, i.e. $\underline{a} = \underline{I}$, $\beta \equiv \mathbf{0}$ in equation (2.1), discretized by the bilinear form $B_1(\cdot,\cdot)$. We further confine ourselves to the symmetric discretizations BR2 and SIPG and postpone the discussion of the non-symmetric interior penalty method to Section 4.3. By C, C_i we will denote generic constants.

4.1. Theoretical tools.

LEMMA 4.1. *There exist constants $C_1, C_2 > 0$ depending on the shape regularity constant of $\mathcal{T}_h$, and C_1 also depending on p, such that for all $\kappa \in \mathcal{T}_h$, and $\phi \in V_h^p$ the following inequalities hold:*

$$|\phi|_{1,\kappa}^2 \le C_1 h^{-2} \|\phi\|_{0,\kappa}^2, \qquad \textit{(local inverse inequality)} \tag{4.1}$$

$$\|\phi\|_{0,\partial\kappa}^2 \le C_2 h^{-1} \|\phi\|_{0,\kappa}^2. \qquad \textit{(local trace inequality)} \tag{4.2}$$

Proof. [32], [31].

LEMMA 4.2 (Poincaré-Friedrichs inequality). *Let $D \subseteq \Omega$ be an open connected polyhedral domain that can be covered by the union of some elements in $\mathcal{T}_h$. Then there exists $C > 0$, depending on D, $\mathcal{T}_h$, such that for all $u \in H^1(\mathcal{T}_h)$*

$$\|u - p\|_{0,D}^2 \le C \operatorname{diam}(D)^2 \left(\sum_{\kappa \in \mathcal{T}_h, \kappa \subset D} |u|_{1,\kappa}^2 + \sum_{e \in \mathcal{E}, e \subset D} h^{-1} \|[\![u]\!]\|_{0,e}^2 \right), \tag{4.3}$$

where $p := \frac{1}{\mu(D)} \int_D u \, d\mathbf{x}$.

Proof. See, for example, [1].

We also recall the following lemma that, in the case of the symmetric interior penalty method, requires a sufficiently large stabilization parameter δ.

LEMMA 4.3 (Coercivity and boundedness [2]). *With respect to the norm* $\|\!|\cdot|\!\|_{dG} : V_h^p \to \mathbb{R}^+$, *defined by*

$$\|\!|v|\!\|_{dG}^2 = \sum_{\kappa\in\mathcal{T}_h} |v|_{1,\kappa}^2 + h^{-1}\sum_{e\in\mathcal{E}} \|[\![v]\!]\|_{0,e}^2, \quad v \in V_h^p,$$

the bilinear form $B_1(\cdot,\cdot)$ *is coercive, i.e. there exists* $C > 0$ *such that*

$$B_1(u,u) \geq C\|\!|u|\!\|_{dG}^2 \quad \forall u \in V_h^p.$$

Further, there exists $C_B > 0$, *such that*

$$B_1(u,v) \leq C_B\|\!|u|\!\|_{dG}\|\!|v|\!\|_{dG} \quad \forall u,v \in V_h^p.$$

Finally, we have a bound for the spectral radius of the stiffness matrix, similar to the standard finite element method.

LEMMA 4.4. *For the spectral radius* $\varrho(\underline{A})$, *where* $\underline{A} \in \mathbb{R}^{n\times n}$ *denotes the stiffness matrix corresponding to* $B_1(\cdot,\cdot)$, *we have*

$$\varrho(\underline{A}) \leq Ch^{d-2} \tag{4.4}$$

with a constant $C > 0$ *depending on the stabilization parameter* δ, *the shape-regularity of* $\mathcal{T}_h$ *and the polynomial degree* p *of* V_h^p.

Proof. The proof relies on a bound for the energy norm $u \mapsto B_1(u,u)$. We have

$$B_1(u,u) = |u|_{1,\mathcal{T}_h}^2 - 2\int_{\Gamma\cup\Gamma_{\text{int}}} [\![u]\!]\{\nabla_h u\}\, ds - s(u) \quad \text{for } u \in V_h^p, \tag{4.5}$$

with a stabilization term $s(u) := \sum_{e\in\mathcal{E}} \delta h^{-1}p^2 \int_e [\![u]\!]^2\, ds$ for the interior penalty method, and $s(u) := \sum_{e\in\mathcal{E}} \eta_e \int_\Omega \left(r_e([\![u]\!])\right)^2\, ds$ for the BR2 method. Using the property of the lifting operator r_e

$$C_1 h^{-1}\|u\|_{0,e}^2 \leq \|r_e(u)\|_{0,\Omega}^2 \leq C_2 h^{-1}\|u\|_{0,e}^2, \quad e \in \mathcal{E},$$

it suffices to consider the SIPG stabilization. Applying Young's inequality together with (4.2) we get

$$\int_{\Gamma\cup\Gamma_{\text{int}}} [\![u]\!]\{\nabla_h u\}\, ds \leq \frac{1}{2h}\int_{\Gamma\cup\Gamma_{\text{int}}} [\![u]\!]^2\, ds + \frac{C}{2}|u|_{1,\mathcal{T}_h}^2. \tag{4.6}$$

Substituting (4.6), (4.2) in equation (4.5) we have

$$B_1(u,u) \leq (1+C)|u|_{1,\mathcal{T}_h}^2 + (1+\delta p^2)\int_{\Gamma\cup\Gamma_{\text{int}}} h^{-1}[\![u]\!]^2\, ds.$$

Applying (4.1), (4.2) once again we get the bound

$$B_1(u,u) \leq Ch^{-2}\|u\|_{0,\mathcal{T}_h}^2,$$

with C depending on δ, p. The bound for the spectral radius of $\underline{A}$ follows then from the equivalence between the discrete norm $\|\cdot\|_2$ and $\|\cdot\|_{0,\mathcal{T}_h}$: Let $\eta \in \mathbb{R}^n$ denote the vector of coefficients for $u = \sum_{i=1}^n \eta_i\phi_i$. Then we have

$$\varrho(\underline{A}) \leq \frac{\eta^T\underline{A}\eta}{\|\eta\|_2^2} \leq \frac{B_1(u,u)}{\|\eta\|_2^2} \leq Ch^{-2}\frac{\|u\|_{0,\mathcal{T}_h}^2}{\|\eta\|_2^2} \leq Ch^{d-2}$$

which concludes the proof of lemma. □

4.2. Convergence estimate. Our V-cycle estimate is an application of the following convergence theorem.

THEOREM 4.5 (Vaněk et al. [35], Guillard et al. [15]). *Assume an agglomeration of the tesselation $\mathcal{T}_h$, together with orthonormal piecewise constant operators $\tilde{I}_{k-1}^k$, $\tilde{I}_k^J$ as defined in* (3.3).

Let the prolongation and restriction smoothers be given by the polynomials (3.4) *or* (3.5) *for the Ritz-Galerkin or the Petrov-Galerkin multigrid, respectively, We assume $\lambda_l := (r+1)^{2(l-J)}\lambda$, $\lambda \geq \varrho(\underline{A})$, and $r = 2$ in the Ritz-Galerkin multigrid, $r = 1$ in the Petrov-Galerkin case.*

We further assume that there exists a constant C_1 independent of the level and linear mappings

$$\underline{\tilde{Q}}_l : \mathbb{R}^{n_J} \to \mathbb{R}^{n_l}, \quad l = 1, \ldots, J, \quad \underline{\tilde{Q}}_J = \underline{I},$$

which satisfy a weak approximation property

$$\|\mathbf{u} - \tilde{I}_l^J \underline{\tilde{Q}}_l \mathbf{u}\|_{\mathbb{R}^{n_J}}^2 \leq C_1^2 \frac{(r+1)^{2l}}{\lambda} \|\mathbf{u}\|_{\underline{A}}^2 \quad \forall \mathbf{u} \in \mathbb{R}^{n_J}, l = 1, \ldots, J-1. \tag{4.7}$$

Further let the smoothers $\underline{R}_k$, $k = 2, \ldots, J$, satisfy the Assumption 1 with constants $C_R > 1$, $\theta \in [0, 2)$. Then

$$\|\underline{A}^{-1}\mathbf{b} - MG(\mathbf{x}, \mathbf{b})\|_{\underline{A}} \leq \left(1 - \frac{1}{C}\right) \|\underline{A}^{-1}\mathbf{b} - \mathbf{x}\|_{\underline{A}}, \tag{4.8}$$

where

$$C = \left[2 + \frac{1}{r+1} C_1 (J-1) + C_R \left(C_1 + C_2 + \frac{1}{r+1} C_1 C_2 (J-1)\right) \sqrt{\frac{\theta}{2-\theta}}\right]^2 \frac{J-1}{2-\theta} \tag{4.9}$$

with $C_2 := \frac{4}{3}$ in the Ritz-Galerkin case and $C_2 := \sqrt{2 + \frac{\pi^2}{3}}$ for the Petrov-Galerkin multigrid. If the multigrid algorithm is used as a preconditioner, then it follows from the Rayleigh quotient characterization of the extreme eigenvalues, that there holds cond $(\underline{B}_{MG}\underline{A}) \leq C$.

Keeping in mind the theoretical tools revisited in Section 4.1 the proof of the weak approximation property (4.7) for the dGFEM is straightforward. We assume that the following geometric properties hold for the agglomeration.

ASSUMPTION 2. *For each composite aggregate $a_{l,i}^J$ of the given agglomeration there exists a ball $\mathcal{U}_i^l \supseteq a_{l,i}^J$, $\mathcal{U}_i^l \subset \mathbb{R}^d$ that fulfils the following properties.*

1. *The overlap of the balls $\{\mathcal{U}_i^l\}_{i \in n_l}$ is locally bounded independent of l, i.e. there exists $N > 0$ s. t. for each $\mathbf{x} \in \Omega$ there are at most N balls containing $\mathbf{x}$.*
2. *With r defined as in Theorem 4.5 it holds*

$$\operatorname{diam}(\mathcal{U}_i^l) \leq C(r+1)^l h. \tag{4.10}$$

The linear mappings $\underline{\tilde{Q}}_l : \mathbb{R}^{n_J} \to \mathbb{R}^{n_l}$, $l = 1, \ldots, J-1$, are then constructed as follows. We define for $u_h \in V_h^p$, represented by $u_h = \sum_{i=1}^{n_J} u_i \phi_i$, $\mathbf{u} = \{u_i\}_{i=1}^{n_J}$:

$$\underline{\tilde{Q}}_k \mathbf{u} = \mathbf{w}^k, \quad \left(\mathbf{w}^k\right)_i = \frac{1}{\mu(\mathcal{U}_i^l)} \int_{\mathcal{U}_i^l} u_h \, d\mathbf{x}, \quad i = 1, \ldots, n_k.$$

Then we have

$$\|\mathbf{u} - \tilde{I}_l^J \underline{\tilde{Q}}_l \mathbf{u}\|_2^2 = \sum_{i=1}^{n_l} \|\mathbf{u} - \tilde{I}_l^J \mathbf{w}^l\|_{\ell^2(a_{l,i}^J)}^2 = \sum_{i=1}^{n_l} \|\mathbf{u} - \left(\mathbf{w}^l\right)_i\|_{\ell^2(a_{l,i}^J)}^2, \tag{4.11}$$

where for a set $T \subseteq \Omega$ the notation $\|\cdot\|_{\ell^2(T)}$ denotes the Euclidean norm restricted to the degrees of freedom contained in T. In the continuous (broken) L^2-norm the expression can be rewritten as

$$\begin{aligned}\|\mathbf{u} - \left(\mathbf{w}^l\right)_i\|_{\ell^2(a_{l,i}^J)}^2 &\leq Ch^{-d}\|u_h - \left(\mathbf{w}^l\right)_i\|_{0,\mathcal{T}_h\cap\mathcal{U}_i^k}^2 \\ &\leq Ch^{-d}\operatorname{diam}\left(\mathcal{U}_i^l\right)^2 \left(\sum_{\kappa\in\mathcal{T}_h,\kappa\subset\mathcal{U}_i^l} |u_h|_{1,\kappa}^2 + \sum_{e\in\mathcal{E},e\subset\mathcal{U}_i^l} h^{-1}\|[\![u_h]\!]\|_{0,e}^2 \right),\end{aligned}$$

where in the last step we have applied the Poincaré-Friedrichs inequality (4.3). Finally, using assumption (4.10) and the coercivity of the bilinear form we get

$$\|\mathbf{u} - \left(\mathbf{w}^l\right)_i\|_{\ell^2(a_{l,i}^J)}^2 \leq Ch^{-d}\left[C(r+1)^l h\right]^2 B_1(u_h,u_h) = Ch^{2-d}(r+1)^{2l}B_1(u_h,u_h).$$

Substituting into equation (4.11) together with (4.4) yields the result

$$\|\mathbf{u} - \tilde{I}_l^J \underline{\tilde{Q}}_l \mathbf{u}\|_2^2 \leq (r+1)^{2l}\frac{1}{\varrho(\underline{A})}\|\mathbf{u}\|_{\underline{A}}^2,$$

which gives the estimate (4.8).

4.3. Preconditioning of non-symmetric discretizations. The convergence result developed in the previous section does not extend to the case of the non-symmetric interior penalty discretization. Nevertheless, error estimates can be established for the NIPG method as well, if we take a GMRES iteration and a multilevel preconditioner for the symmetric part

$$B_0(u,v) := \int_\Omega \nabla_h u \cdot \nabla_h v \; d\mathbf{x} + \int_{\Gamma\cup\Gamma_{\text{int}}} [\![u]\!]\cdot[\![v]\!] \; ds,$$

corresponding to the stiffness matrix $\underline{A}_0 = \frac{1}{2}\left(\underline{A}_{NIPG} + \underline{A}_{NIPG}^T\right)$.

Let $c_1, c_2 > 0$ be defined as

$$c_1 = \lambda_{\min}\left(\underline{B}_{MG}\underline{A}_0\right), \quad c_2 = \|\underline{B}_{MG}\underline{A}_{\text{NIPG}}\|_{\underline{A}_0}.$$

Then, following the theory for minimal residual Krylov methods [10], the norm of the residual $\mathbf{r}^k$ after k steps of the GMRES method is bounded by

$$\|\mathbf{r}_k\|_{\underline{A}_0} \leq \left(1 - \frac{c_1^2}{c_2^2}\right)^{\frac{k}{2}} \|\mathbf{r}_0\|_{\underline{A}_0}.$$

Note that this result is valid for the GMRES method formulated wrt. the energy norm $\|\cdot\|_{\underline{A}_0}$, but it is of importance for the Euclidean inner product iteration as well, since there exist constants C_1, C_2, such that

$$C_1 h^d \|\mathbf{x}\|_2^2 \leq \|\mathbf{x}\|_{\underline{A}_0}^2 \leq C_2 h^{d-2}\|\mathbf{x}\|_2^2.$$

Thus the iteration count is increased by a factor of order $\ln h$.

Usually, convergence bounds for preconditioners applied to non-symmetric forms are obtained using a lemma by Schatz and a duality argument, cf. [33]. This technique cannot be applied to the NIPG bilinear form which is not adjoint consistent [2]. Hence, in the following we directly derive estimates for the constants c_1, c_2. First we need the following result.

LEMMA 4.6. *There exists a constant $\gamma > 0$, s. t.*

$$B_{NIPG}(v,w) \leq \gamma \left[B_0(v,v)\right]^{\frac{1}{2}} \left[B_0(w,w)\right]^{\frac{1}{2}} \quad \forall v \in V_h^p.$$

Proof. See, e. g., [1].

Now the following lemma is a consequence of Theorem 4.5:

LEMMA 4.7. *Let the non-symmetric interior penalty discretization be solved with the GMRES algorithm. We further assume that the multilevel preconditioner described in Section 3 is formulated for the symmetric part $\underline{A}_0$ of the stiffness matrix. Then there holds*

$$c_1 := \lambda_{min}\left(\underline{B}_{MG}\underline{A}_0\right) \geq \frac{1}{C} \quad \textit{and} \quad c_2 := \|\underline{B}_{MG}\underline{A}_{NIPG}\|_{\underline{A}_0} \leq \gamma$$

with the constant C given in (4.9).

Proof. The lower bound follows directly from Theorem 4.5, since $\underline{A}_0$ is symmetric, and all assumptions hold due to the spectral equivalence between $B_{\mathrm{SIPG}}(\cdot,\cdot)$ and $B_0(\cdot,\cdot)$, cf. Lemmata 4.3 and 4.6.

For the upper estimate, we have

$$\begin{aligned}
\|\underline{B}_{MG}\underline{A}_{\mathrm{NIPG}}\mathbf{x}\|_{\underline{A}_0}^2 &= (\underline{A}_0\underline{B}_{MG}\underline{A}_{\mathrm{NIPG}}\mathbf{x}, \underline{B}_{MG}\underline{A}_{\mathrm{NIPG}}\mathbf{x}) \\
&= (\underline{A}_{\mathrm{NIPG}}\mathbf{x}, \underline{B}_{MG}\underline{A}_0\underline{B}_{MG}\underline{A}_{\mathrm{NIPG}}\mathbf{x}) \\
&\leq \gamma \|(\underline{B}_{MG}\underline{A}_0\underline{B}_{MG})\underline{A}_{\mathrm{NIPG}}\mathbf{x}\|_{\underline{A}_0} \|\mathbf{x}\|_{\underline{A}_0} \\
&\leq \gamma \|\underline{B}_{MG}\underline{A}_{\mathrm{NIPG}}\mathbf{x}\|_{\underline{A}_0} \|\mathbf{x}\|_{\underline{A}_0},
\end{aligned}$$

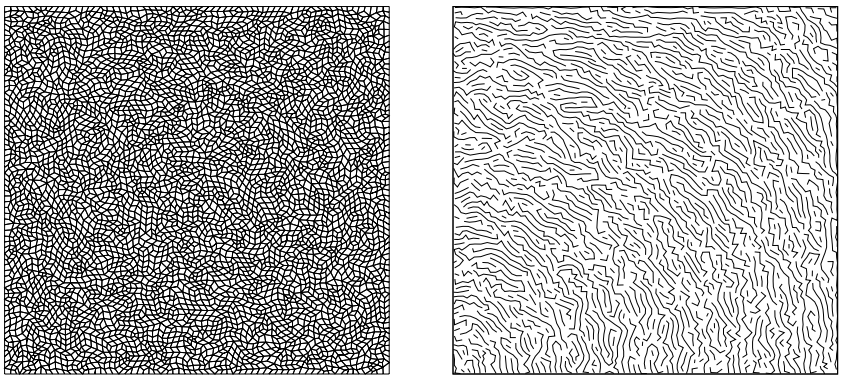
where we used $\sigma(\underline{I} - \underline{B}_{MG}\underline{A}_0) \subset [0, 1 - \frac{1}{C}]$. Thus the norm of the operator is bounded by γ. ∎

5. Numerical examples. In this section we demonstrate the feasibility of the smoothed aggregation multigrid for the dG method. Numerical results have been obtained using the `PADGE` code [18] with the support of the `deal.II` libraries [3], Argonne's PETSc library and the LAPACK linear algebra package.

5.1. Setting. With the exception of the test case including the NACA 0012 airfoil, that is discussed in Section 5.4, we restrict ourselves to the test domain $\Omega = (0,1)^2$. Two sequences of tesselations are considered for scalability experiments: The first sequence, $\{\mathcal{T}_i^s\}_{i=1}^5$, consists of nested quadrilateral meshes constructed by global uniform refinement of the unit square. The corresponding grid sizes are $h = \frac{1}{32}, \frac{1}{64}, \ldots, \frac{1}{512}$. The second mesh sequence, $\{\mathcal{T}_i^u\}_{i=1}^5$, is a collection of unstructured, isotropic, non-nested tesselations with $N = \{1335, 5316, 21104, 84269, 337076\}$ elements. The second-coarsest one, $\mathcal{T}_2^u$, is depicted in Figure 5.1(a).

The algorithmic scalability is measured by the average and the asymptotic residual reduction, denoted by r_{avg} and $r_{avg,k}$. After N iterations of the V-cycle evaluated in the Euclidean norm we have

$$r_{avg} := \left(\frac{\|r^N\|_2}{\|r^0\|_2}\right)^{\frac{1}{N}}, \quad r_{avg,k} := \left(\frac{\|r^N\|_2}{\|r^{N-k}\|_2}\right)^{\frac{1}{k}}, \quad k < N.$$



(a) Example mesh, 5316 elements. (b) Constructed lines.

FIGURE 5.1. *Result of the line creation heuristic. Advection vector* $\beta(\mathbf{x}) = (-x_2, x_1)^T$*, the lines formed by elements are represented by curves through the centers.*

Additionally, in view of Theorem 4.5 we also approximate the reduction of the exact error in the energy norm $\|\underline{A}^{-1}\mathbf{b} - \mathbf{x}^N\|_{\underline{A}}$.

5.2. Poisson equation. Since we are interested in the ideal scalability of the algorithm, we first consider a simple experiment with Poisson's equation having homogeneous Dirichlet boundary conditions and a constant source term $f \equiv 1$. Theorem 4.5 predicts a condition number independent from the number of unknowns for the linear algebraic system preconditioned by the smoothed aggregation algorithm.

As a simple smoothing iteration for this problem we apply 2 pre- and post-smoothing steps of the pointwise, symmetric Gauß-Seidel method, abbreviated as V(2,2). The spectral radii $\varrho(\underline{A}_k)$, $k = 2, \ldots, J$, required for the construction of the smoothing polynomials, are approximated by a few steps of the Lanczos algorithm. In all examples, a Krylov type iterative method has been employed together with the V-cycle preconditioner. We used the so-called *flexible* variant [34] of the GMRES method minimizing the ℓ^2-norm of the residual due to the variable behavior of the multigrid preconditioner during the Krylov method iteration. The number of levels varied from 5 to 9, such that the coarsest level consisted of less than 16 unknowns. Instead of a maximum number of iterations an absolute tolerance of $\|\underline{A}\mathbf{x}^n - \mathbf{b}\| \leq 10^{-10}$ for the Euclidean norm of the residual (coefficient) vector was chosen as an abort criterion.

In Figure 5.2 the asymptotic residual reduction $r_{avg,5}$ is shown for the structured and unstructured mesh sequences. The values correspond to a piecewise bilinear SIPG discretization ($\delta := 10$) and the Bassi-Rebay scheme. They are plotted over the square-root of elements, which is roughly the grid size h for the unstructured tesselations. Both, the symmetric and the non-symmetric multigrid variant scale optimally with h on the structured mesh sequence, and the average residual reduction seems to be bounded. Table 5.1 reflects the same behavior for the error reduction that is measured in the energy norm. Finally, the theoretical results are supported by the condition number estimates in Table 5.2, whereas the piecewise constant transfer operators cause the multilevel algorithm to deteriorate.

The Ritz-Galerkin multigrid yields better results than the Petrov-Galerkin approach, however, we do not take into account the increased computational effort,

which is caused by the higher density of the coarse level matrices. The additional pre-smoothing step for the Petrov-Galerkin multigrid has been omitted in the examples, as it did not seem to affect the convergence behavior of the method.

Results are less satisfactory for the unstructured mesh sequence, at least there is no clear evidence for an asymptotic behavior. This may be caused by the irregular shape of the agglomerates, which can already be seen in Figure 3.1. Further, the convergence result stated in Section 4.3 for the non-symmetric interior penalty discretization was verified for Poisson's equation. The experimental rates of convergence given in Table 5.3 are qualitatively similar to the results for the symmetric dGFEM variants. However, the residual reduction is generally slower in absolute values.

Furthermore, the multilevel method can be accelerated by a damped coarse grid correction step

$$\mathbf{x}^l \leftarrow \mathbf{x}^l + \alpha I_{k-1}^k \mathbf{x}^{l-1}$$

with a factor $\alpha \in (0, 1]$. The effect of this *overcorrection* is shown in Figure 5.3(a) for the structured mesh sequence. Here, the scaling α of the coarse level operators has been determined experimentally and yields an improvement of the convergence rate r_{avg} of about $\Delta r_{avg} \approx 0.1$.

Now let us concentrate on the computational complexity of the method. Following the arguments in Section 3.2, the dimensions n_i, $i = 1, \ldots, J$, of the level subspaces have to increase geometrically with i, which is determined by the agglomeration strategy. Further, the amount of work per cycle, except for the coarsest level, must behave like Cn_i. This yields an operation count like $\mathcal{O}(n_J)$ for the whole V-cycle.

This is investigated in Table 5.4. It compares the nonzeros, denoted by #nnz, which are contained in the smoothed aggregation level matrices to the case of piecewise constant prolongation and restriction. The additional fill-in by the smoothed aggregation approach is significant, especially for piecewise constant finite elements, though much less severe for the Petrov-Galerkin approach than for the Ritz-Galerkin multigrid. Concerning the computational work the Ritz-Galerkin multigrid is suboptimal due to the increased matrix fill-in. However, in terms of memory requirements, the fine level matrix outweighs the hierarchy of coarse matrices by far and the two approaches are more or less comparable.

Further, we examine the acceleration from the coarse levels of the hierarchy. This test is performed by varying the number of grid levels for the elliptic model problem, where instead of an exact solution on the coarsest level only pre- and post-smoothing is done, analogous to the other levels in the hierarchy. In Figure 5.3(b) it is shown that the efficiency of the truncated V-cycle is no longer linear, but it improves with an increased number of levels until it resembles the behavior of the complete V-cycle.

5.3. Linear convection-diffusion problem. We now return to the second order differential equation (2.1) and consider the convection-diffusion problem, which is defined by the trivial diffusion matrix $\underline{a} \equiv \mu \underline{I}$ and the vector field $\beta(\mathbf{x}) = \gamma\,(-x_2, x_1)^T$. We set $\mu = 10^{-3}$ and let $\gamma \in \mathbb{R}^+$ denote a variable scaling factor of the rigid rotation. The equation is complemented with Dirichlet boundary conditions for $\Gamma_\mathrm{D} = \{\mathbf{x} \in \Gamma : (x_1 = 0) \vee (x_2 = 1)\}$, while on the remaining part, we apply a Neumann boundary condition, $\nabla u \cdot \mathbf{n} = g_\mathrm{N} \equiv 0$. The data g_D on the Dirichlet boundary

Γ_{D}, which contains the inflow part Γ_-, is given by

$$g_{\mathrm{D}}(x_1, x_2) := \begin{cases} 5(x_1 - 0.2) & \text{for} \quad 0.2 < x_1 \leq 0.4, x_2 = 0, \\ 1 & 0.4 < x_1 \leq 0.6, x_2 = 0, \\ 1 - 5(x_1 - 0.6) & 0.6 < x_1 \leq 0.8, x_2 = 0, \\ 0 & \text{elsewhere.} \end{cases}$$

We are interested in the behavior of the multigrid algorithm in the case of varying advection skew to the mesh. Therefore, the average residual reduction r_{avg} is computed over a range of values for the Péclet number $\mathrm{Pe} := \|\beta\|_\infty h \mu^{-1}$, where the maximum speed of advection occurring in the domain of interest is $\|\beta\|_\infty = \gamma$. The results shown in the Figures 5.4, 5.5 were computed for the structured mesh sequence $\{\mathcal{T}_i^s\}_{i=1}^5$ with bilinear finite elements. The algorithm was chosen to be a GMRES accelerated multigrid with a V(1,1) (non-symmetric) line-implicit Gauß-Seidel smoothing iteration.

We discuss the results for Ritz-Galerkin multigrid first. The set of the studied test cases can be divided into three different flow regimes, see Figure 5.4. In the first one, approximately $\mathrm{Pe} \in [10^{-4}; 10^{-1}]$, the elliptic term stays predominant. As it can be expected from the h-optimality tests in Section 5.2 the average residual reduction is approximately the same for all meshes.

On the other hand, in the case of predominant convection, $\mathrm{Pe} > 100$, optimality of the algorithm is not fulfilled, i. e. the convergence rate increases with the number of unknowns. In Figure 5.5, which will be discussed in more detail below, the important role of the smoothing iteration for the hyperbolic term becomes evident. The behavior of the multigrid iteration improves significantly when employing a directionally implicit smoothing iteration, since the Gauß-Seidel method degrades into a (nearly) explicit solver. Another observation in Figure 5.4 is the fact that both approaches, the unsmoothed aggregation approximated by the unsplit operator approach and the smoothed aggregation variant, yield comparable residual reductions in the hyperbolic regime. The smoothing polynomial clearly does not play a role here.

The effect of the operator splitting $\underline{\underline{A}} = \underline{\underline{A}}^v + \underline{\underline{A}}^c$ with a separate treatment of the convective and the diffusive operator, which was discussed in Section 3.5, is visible in the flow regime $\mathrm{Pe} \in [10^{-1}; 100]$, where the diffusive and convective components balance each other. Compared to the smoothed aggregation algorithm without operator splitting, the average residual reduction is improved for all grid sizes.

At this point, we illustrate another interesting modification of the proposed algorithm. As mentioned before, there is some freedom in distributing the unknowns on the first agglomerated level. Without destroying the algorithmic optimality, the level $J - 1$ can be constructed as a trivial agglomeration, i. e. with $n_\kappa = n_t$ agglomerates. We receive a method similar to the two-level iteration proposed by Dobrev et al. [9]. At the expense of a larger memory consumption, the average convergence rate improves significantly, approximately by $\Delta r_{avg} = 0.1$.

Finally, we discuss the Petrov-Galerkin multigrid. Basically, the observations made for the Ritz-Galerkin case also hold true for the Petrov-Galerkin smoothed aggregation, and details were omitted for brevity. Similar to the elliptic model problem, however, the average reduction r_{avg} is slightly increased. In the regime of dominant convection, the smoothing polynomials have no effect and the residual reductions for the Ritz-Galerkin and the Petrov-Galerkin multigrid coincide.

As we are considering the sequence of nested, structured tesselations for this test problem, the smoothed aggregation multigrid can directly be compared to the

classical geometric multigrid, where the hierarchy of levels stems from the previous refinement process. We choose the same problem parameters as before, but use an SIPG stabilization with $\delta := 10$. For the interpolation operators I_k^{k-1}, I_{k-1}^k, we use the injection $\iota : V_h^p \hookrightarrow V_H^p$, where $H := 2h$, and the bilinear form $B(\cdot,\cdot)$ is discretized with bilinear finite elements on each level. The algorithm uses 2 pre- and post-smoothing steps of the block Gauß-Seidel method and is described in [14].

Results are shown in Figure 5.5. In the regime of predominant diffusion the geometric algorithm is superior to the algebraic approach. However, this changes with an increasing convection term, and the algorithm performs worse than the algebraic multigrid for (nearly) hyperbolic problems. This is likely to be caused by a poor smoothing iteration. The same can be observed for the smoothed aggregation multigrid, where an element-wise block Gauß-Seidel method was used instead of the line-implicit scheme. Additionally, a suitable renumbering of the unknowns has a great effect on the convergence rate. Overall, we should point out that this comparison does not take into account the absolute computational effort, where the geometric multigrid is clearly superior due to its simpler algorithmic nature.

5.4. Linearized Navier-Stokes equations. As an example for nonlinear convection-diffusion, we consider a Newton-multigrid approach for the Navier-Stokes equations. For laminar compressible flow the conservative state is given by $\mathbf{u} = [\rho, \rho\mathbf{v}, \rho E]^T$ with ρ denoting the density, $\mathbf{v}$ the velocity vector and E the specific total energy. The Navier-Stokes equations are formulated in the Cartesian coordinate system as

$$\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathcal{F}^c(\mathbf{u}) - \mathcal{F}^v(\mathbf{u}, \nabla\mathbf{u})) = \mathbf{0} \text{ in } \Omega \subset \mathbb{R}^d \tag{5.1}$$

with convective and viscous fluxes $\mathcal{F}^c(\mathbf{u}) = (\mathbf{f}_1^c, \ldots, \mathbf{f}_d^c)^T$, $\mathcal{F}^v(\mathbf{u}, \nabla\mathbf{u}) = (\mathbf{f}_1^v, \ldots, \mathbf{f}_d^v)^T$. These are given by

$$\mathbf{f}_s^c(\mathbf{u}) = \left(\rho v_s, \rho v_1 v_s + \delta_{1s} p, \ldots, \rho v_d v_s + \delta_{ds} p, (\rho E + p) v_s\right)^T, s = 1, \ldots, d,$$

with $p = (\gamma - 1)\left[\rho E - \frac{1}{2}\|\mathbf{v}\|_2^2\right]$ denoting the pressure, the Poisson adiabatic constant $\gamma = 1.4$ (dry air), and

$$\mathbf{f}_s^v(\mathbf{u}, \nabla\mathbf{u}) = \left(0, \tau_{s1}, \ldots, \tau_{sd}, \sum_{i=1}^{d} \tau_{si} v_i + \mathcal{K}\frac{\partial T}{\partial x_s}\right)^T, s = 1, \ldots, d.$$

Here, T denotes the temperature function, $\mathcal{K}$ the heat conduction coefficient and τ is the viscous stress tensor. The Navier-Stokes equations are complemented by suitable Dirichlet and Neumann boundary conditions

$$\begin{cases} D\left(\mathbf{u}\big|_\Gamma\right) = 0 \\ N\left(\mathcal{F}^v(\mathbf{u}, \nabla\mathbf{u}) \cdot \mathbf{n}\big|_\Gamma\right) = 0, \quad \mathbf{x} \text{ in } \Omega, \end{cases}$$

see, for example, the monograph by Feistauer [11] for details. The dG semilinear form is then derived similar to the scalar case via reformulation as a first order system and the substitution of stabilizing numerical fluxes [4, 19]. In our experiments the BR2 stabilization and bilinear finite elements were employed. For the numerical flux function we used the Vijayasundaram scheme [11].

Since we are interested in steady solutions, the variable $t > 0$ plays the role of a pseudo-time variable, and the temporal accuracy is irrelevant. We linearize (5.1) with a semi-implicit Euler scheme

$$\left[\frac{1}{\Delta t}\underline{M} + \underline{J}(\mathbf{u}^n)\right]\mathbf{d} = -\nabla\cdot(\mathcal{F}^c(\mathbf{u}^n) - \mathcal{F}^v(\mathbf{u}^n, \nabla\mathbf{u}^n)) =: -\mathcal{N}(\mathbf{u}^n), \quad \mathbf{d} = \mathbf{u}^{n+1} - \mathbf{u}^n, \tag{5.2}$$

where $\underline{J}(\mathbf{u}^n) := \frac{\partial\mathcal{N}}{\partial\mathbf{u}}(\mathbf{u}^n)$ denotes the Jacobian, $\underline{M} = \left\{(\phi_i, \phi_j)_{L^2(\Omega)}\right\}_{i,j}^n$ the mass matrix, and Δt is a suitable time step. For $\Delta t \to \infty$ we recover Newton's method with $\mathbf{d}$ as a search direction. Thus the artificial time can be viewed as a globalization strategy. In the example, we employed a local time step strategy controlled by a fixed global CFL number, $CFL = 10^2$, that is deactivated after 6 nonlinear steps.

The discrete linearized problem (5.2) exhibits a similar structure $\underline{J} = \underline{A}^v + \underline{A}^c$ as the convection-diffusion model case, however, it now corresponds to a vector-valued problem. Treating all degrees of freedom associated with a shape function as a block, usually identified with a node, the multigrid approach described in Section 3 can be applied to the linearized Navier-Stokes equations. The recomputation of the interpolation and restriction operator in each step of the outer nonlinear iteration is avoided by constructing the smoothed transfer operators with respect to the discrete Laplacian [22].

We restrict ourselves to a basic two-dimensional problem geometry. The test case under consideration is the well-known symmetric NACA 0012 airfoil with $\alpha = 0°$ angle of attack. We pose adiabatic no-slip boundary conditions at the wall and characteristic far-field boundary conditions. The flow is characterized by the quantities

$$\text{Ma} = 0.5, \quad \text{Re} = 5000,$$

which also determines the initial freestream state $\mathbf{u}^0$. The (structured) tesselations consist of 768, 3072 and 12288 quadrilaterals, respectively. Figure 5.6(a) shows a Mach number contour plot of the discrete solution.

In addition to the flexible GMRES algorithm, the BiCGStab [34] algorithm was examined, which requires less memory resources and usually exhibits faster convergence, but which is also less stable. From an implementational point of view the Newton-multigrid approach does not require the storage of the fine level stiffness matrix $\underline{A}_J$. Matrices are explicitly stored only for the preconditioner. The necessary matrix-vector products during the Krylov iteration can be computed by suitable finite difference quotients.

As outlined in Section 3.4 the coupling criterion of the line-implicit smoother is based on the momentum components $\rho\mathbf{v}$ of the flow state iterate $\mathbf{u}^k$. Figure 5.6(b) shows an illustration of the created lines, where the viscous effects have been well captured by the coupling criterion. To improve the stability of the smoothing iteration the line-implicit Gauß-Seidel scheme was extended to a multistage method. For this widely used approach [30] the smoothing iteration matrix $\underline{R}$ is replaced by a polynomial in $\underline{R}$ with optimized coefficients. Using the notation of an explicit Runge-Kutta method applied to the coefficient vector $\mathbf{x}^k$ we have

$$\mathbf{k}_0 = \mathbf{x}^k, \quad \mathbf{k}_j = \mathbf{x}^k + \alpha_j\underline{R}(\mathbf{b} - \underline{A}\mathbf{k}_{j-1}),\ j = 1, \ldots, s; \quad \mathbf{x}^k \leftarrow \mathbf{k}_s.$$

The coefficients $\{\alpha_j\}_{j=1,\ldots,s} := \{0.2075, 0.5915, 1\}$ have been chosen according to [30].

Figure 5.7 shows some results for the Newton-multigrid algorithm preconditioned by a V(1,1)-cycle (Ritz-Galerkin). Three representative subproblems during the nonlinear convergence process, cf. A, B, C in Figure 5.7(a), have been plotted in Figure 5.7(b). The multilevel method performs significantly better in the presence of a pseudo-time augmentation of the system matrix $\underline{J}(\mathbf{u})$. Considering the first nonlinear iteration, Table 5.5 lists the number of V-cycles required to reach a reduction of the residual norm by a factor of 10^{-10}. The applied Krylov type iteration plays an important role and the BiCGStab iteration requires about half of the steps for the GMRES method. For this method, however, the residual norm is no longer monotonically decreasing. Furthermore, due to the convective term and the simplifications made for the inter-grid transfer operators, the convergence rate of the algorithm is no longer independent of the grid size h.

We make some final remarks on the computations. First of all, the employed agglomeration heuristic is likely to be inappropriate for the structured meshes. The high aspect ratio of the elements in the wake region leads to deformed agglomerates, violating Assumption 2. A directional coarsening strategy [22, 30] would help improving the behavior of the algorithm. This can be a point of future study.

Second, the problem setting itself is rather simple and generalization to numerical simulation of transonic or supersonic flow, or to the purely hyperbolic system of the Euler equations requires further analysis. The smoothing iteration, which has a great effect on the convergence of the method, must be carefully adjusted.

Third, it is unrealistic for practical applications to apply an absolute-value abort criterion to the linear subproblems. Usually, a criterion related to the reduction of the residual norm relative to $\|\mathbf{r}^0\|_2$ will make sense. Similarly, the fixed CFL number would be replaced by an adaptive time stepping strategy.

6. Conclusion. The concept of smoothed aggregation multigrid developed for the conforming finite element method on unstructured grids has been successfully applied to the discontinuous Galerkin discretization. The scalability of the algorithm has been analyzed theoretically and results were supported by a number of numerical experiments. For the elliptic model problem the rate of convergence has shown to be independent of the number of unknowns. When considering equations of convection-diffusion type the intuitive nature of the aggregation multigrid lends itself to a combination with agglomeration approaches for the hyperbolic part of the problem operator. In this context, additional strategies from geometric multigrid, such as line-implicit smoothers and multistaging were considered.

However, there are still some open questions. To our best knowledge, convergence theory for the smoothed aggregation approach is restricted only to the model case of scalar elliptic problems. Moreover, the results are qualitative in nature, i. e. we typically obtain results with some generic constants. A proof of scalability and robustness for more general forms would be desirable.

Apart from theory, the implementation of the smoothed aggregation multigrid exhibits some disadvantages. As for any global solution method, the algorithm is difficult to implement in parallel. Most important, the approach operates on linear systems containing considerable fill-in. This issue may become prohibitive in terms of memory, however, it hopefully will be mitigated by future hardware developments.

Several problems that remain open will be the topic of our further work. The multigrid algorithm needs to be combined with a p-multilevel method or some related domain decomposition approach to yield a hierarchical solver for high order numerical systems. Furthermore, the treatment of the convective terms needs improvement.

		$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
point Gauß-Seidel	Ritz-Galerkin	0.329	0.345	0.348	0.348	0.349
	Petrov-Galerkin	0.499	0.534	0.534	0.534	0.543
	Ritz-Galerkin, $p = 0$	0.044	0.061	0.073	0.076	0.079
element Jacobi	Ritz-Galerkin	0.477	0.481	0.480	0.472	0.470
	Petrov-Galerkin	0.580	0.585	0.586	0.575	0.558

		$\mathcal{T}_1^u$	$\mathcal{T}_2^u$	$\mathcal{T}_3^u$	$\mathcal{T}_4^u$	$\mathcal{T}_5^u$
point G. S.	Ritz-Galerkin	0.437	0.458	0.502	0.500	0.593
	Petrov-Galerkin	0.580	0.611	0.651	0.661	0.689

TABLE 5.1

Average error reduction with respect to the energy norm $\|\cdot\|_{\underline{A}}$. Results are shown for the structured (top) and unstructured mesh sequence. BR2 discretization.

	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
BR2, point G. S.	4.551	4.652	4.719	4.749	4.779
BR2, element Jacobi	10.363	10.421	10.441	10.452	10.451
SIPG, point G. S.	6.359	6.472	6.548	6.589	6.627
BR2, simple prolongation $\tilde{I}_{k-1}^k$	20.620	33.648	93.219	79.984	246.719

TABLE 5.2

Condition number estimate $\mathrm{cond}\,(\underline{B}_{MG}\underline{A})$.

	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$	$\mathcal{T}_1^u$	$\mathcal{T}_2^u$	$\mathcal{T}_3^u$	$\mathcal{T}_4^u$	$\mathcal{T}_5^u$
$r_{avg,5}$	0.409	0.438	0.420	0.430	0.401	0.491	0.560	0.552	0.582	0.857
r_{avg}	0.437	0.454	0.459	0.462	0.463	0.516	0.575	0.579	0.605	0.782

TABLE 5.3

Average and asymptotic residual reduction for the solution of Poisson's problem on a sequence of refined grids. NIPG discretization.

$p = 0$	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
Petrov-Galerkin	1.158	1.179	1.189	1.195	1.197
Ritz-Galerkin	1.475	1.634	1.754	1.853	1.927
$p = 1$	$\mathcal{T}_1^s$	$\mathcal{T}_2^s$	$\mathcal{T}_3^s$	$\mathcal{T}_4^s$	$\mathcal{T}_5^s$
Petrov-Galerkin	1.012	1.014	1.015	1.016	1.016
Ritz-Galerkin	1.037	1.051	1.061	1.069	1.076

TABLE 5.4

Ratio $\frac{1}{a_0}\sum_{k=1}^{J}\#nnz(\underline{A}_k)$ for the smoothed aggregation coarse operators, SIPG discretization. By a_0 we denote the number of nonzeros when applying piecewise constant smoothing polynomials.

	768 elts.	3072 elts.	12288 elts.
GMRES	22 ($r_{avg} = 0.346$)	25 (0.391)	27 (0.416)
BiCGStab	12	14	14

TABLE 5.5

NACA 0012 test case. Number of V-cycles required to reach a relative tolerance $TOL < 10^{-10}$ for Ritz-Galerkin multigrid, first linear subproblem.

This involves the formulation of the transfer operators as well as the development of better smoothing methods in order to deal with the linearized flow equations.

Acknowledgements. F. Prill and R. Hartmann (e-mail: `ralf.hartmann@dlr.de`) acknowledge the partial financial support of both the President's Initiative and Networking Fund of the Helmholtz Association of German Research Centres and the European project ADIGMA.

The research of M. Lukáčová (e-mail: `lukacova@tu-harburg.de`) was supported partially by the Deutsche Forschungsgemeinschaft under the grant LU 1470/1-1 and by the European Graduate School Differential Equations with Applications in Science and Engineering (DEASE), MEST-CT-2005-021122. The authors gratefully acknowledge these supports.

REFERENCES

[1] P. F. Antonietti, *Domain Decomposition, Spectral Correctness and Numerical Testing of Discontinuous Galerkin Methods*, PhD thesis, Universita degli Studi di Pavia, Dipartimento di Matematica, 2006.

[2] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, *Unified analysis of discontinuous Galerkin methods for elliptic problems*, SIAM J. Numer. Anal., 39 (2002), pp. 1749–1779.

[3] W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II — a general-purpose object-oriented finite element library*, ACM Trans. Math. Softw., 33 (2007).

[4] F. Bassi, A. Crivellini, S. Rebay, and M. Savini, *Discontinuous Galerkin solution of the Reynolds-averaged Navier–Stokes and $k-\omega$ turbulence model equations.*, Comput. Fluids, 34 (2005), pp. 507–540.

[5] J. H. Bramble, *Multigrid methods.*, vol. 294 of Pitman Research Notes in Mathematics Series, Longman Scientific & Technical, Harlow, 1993.

[6] J. H. Bramble and J. E. Pasciak, *The analysis of smoothers for multigrid algorithms.*, Math. Comput., 58 (1992), pp. 467–488.

[7] S. C. Brenner and J. Zhao, *Convergence of multigrid algorithms for interior penalty methods.*, ANACM, Appl. Numer. Anal. Comput. Math., 2 (2005), pp. 3–18.

[8] T. Chan and W. Wan, *Robust multigrid methods for elliptic linear systems*, 1999.

[9] V. A. Dobrev, R. D. Lazarov, P. S. Vassilevski, and L. T. Zikatanov, *Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations*, Numerical Linear Algebra with Applications, 13 (2006), pp. 753–770.

[10] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations.*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.

[11] M. Feistauer, J. Felcman, and I. Straškraba, *Mathematical and computational methods for compressible flow.*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2003.

[12] X. Feng and O. A. Karakashian, *Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems.*, SIAM J. Numer. Anal., 39 (2001), pp. 1343–1365.

[13] K. J. Fidkowski, T. A. Oliver, J. Lu, and D. L. Darmofal, *p-Multigrid solution of high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations*, Journal of Computational Physics, 207 (2005), pp. 92–113.

[14] J. Gopalakrishnan and G. Kanschat, *A multilevel discontinuous Galerkin method.*, Numer. Math., 95 (2003), pp. 527–550.

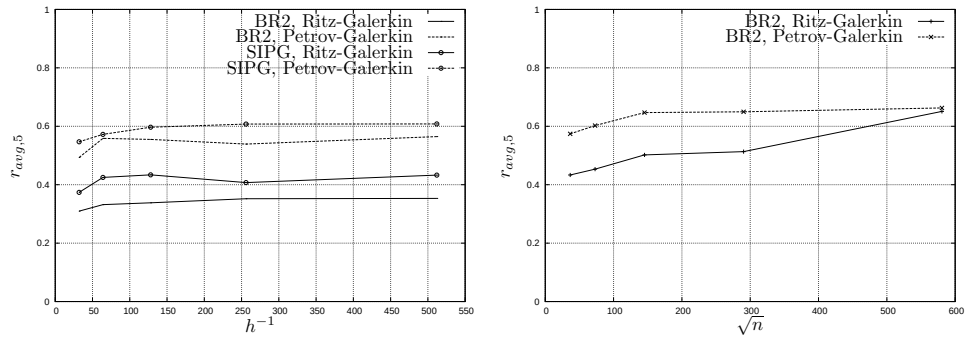
[15] H. Guillard, A. Janka, and P. Vaněk, *Analysis of an Algebraic Petrov-Galerkin Smoothed Aggregation Multigrid Method*, Journal of Applied Numerical Mathematics (accepted), (2006).

[16] H. Guillard and P. Vaněk, *An aggregation multigrid solver for convection-diffusion problems on unstructured meshes.*, Tech. Rep. UCD-CCM-130, University of Colorado in Denver, 1998.

[17] W. Hackbusch, *Multi-Grid Methods and Applications.*, Springer Series in Computational Mathematics, Springer, Berlin, 1985.

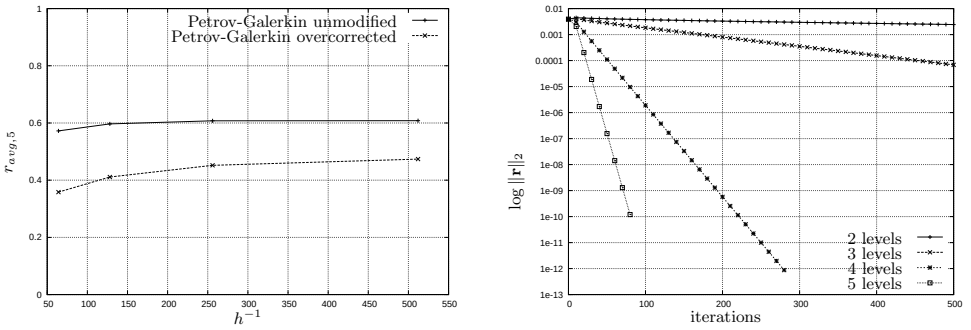
[18] R. Hartmann, J. Held, T. Leicht, and F. Prill, `PADGE`*, Parallel Adaptive Discontinuous Galerkin Environment.* Technical reference, DLR, Braunschweig, 2008.

[19] R. Hartmann and P. Houston, *Symmetric Interior Penalty DG Methods for the Compressible Navier-Stokes Equations*, Int. J. Num. Anal. Model., 3 (2006), pp. 1–20.
[20] P. Hemker, *On the order of prolongations and restrictions in multigrid procedures.*, J. Comput. Appl. Math., 32 (1990), pp. 423–429.
[21] P. Hemker, W. Hoffmann, and M. van Raalte, *Fourier two-level analysis for discontinuous Galerkin discretization with linear elements*, Numerical Linear Algebra with Applications, 11 (2004), pp. 473–491.
[22] A. Janka, *Multigrid Methods for Compressible Laminar Flow*, PhD thesis, INRIA Nice, 2002.
[23] J. E. Jones and P. S. Vassilevski, *AMGe based on element agglomeration*, SIAM J. Sci. Comput., 23 (2001), pp. 109–133.
[24] J. Kraus and J. Synka, *An agglomeration-based multilevel-topology concept with application to 3D-FE meshes*, tech. rep., RICAM, 2004.
[25] J. Kraus and S. Tomar, *Multilevel preconditioning of elliptic problems discretized by a class of discontinuous Galerkin methods.*, tech. rep., Johann Radon Institute for Computational and Applied Mathematics (RICAM), 2006.
[26] C. Lasser and A. Toselli, *Overlapping preconditioners for discontinuous Galerkin approximations of second order problems*, in Thirteenth International Conference on Domain Decomposition Methods, N. Debit, M. Garbey, R. Hoppe, J. Périaux, D. Keyes, and Y. Kuznetsov, eds., 2001.
[27] C. Lasser and A. Toselli, *An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems*, Math. Comput, 72 (2003), pp. 1215–1238.
[28] D. Mavriplis and V. Venkatakrishnan, *Agglomeration multigrid for two-dimensional viscous flows.*, Comput. Fluids, 24 (1995), pp. 553–570.
[29] D. J. Mavriplis, *Multigrid techniques for unstructured meshes*, in ICASE Report No. 95-27, Lecture notes for 26th CFD Lecture Series of von Karman Institute, AGARD Publication, April 1995.
[30] T. O. Okusanya, *Algebraic multigrid for stabilized finite element discretizations of the Navier-Stokes equations*, PhD thesis, Massachusetts Institute of Technology. Dept. of Aeronautics and Astronautics, 2002.
[31] S. Prudhomme, F. Pascal, J. T. Oden, and A. Romkes, *Review of A Priori Error Estimation for Discontinuous Galerkin Methods*, tech. rep., TICAM, Austin, 2000.
[32] A. Quarteroni and A. Valli, *Numerical approximation of partial differential equations.*, no. 23 in Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1994.
[33] A. Toselli and O. Widlund, *Domain decomposition methods – algorithms and theory*, vol. 34 of Springer Series in Computational Mathematics, Springer, Berlin, 2005.
[34] H. A. van der Vorst, *Iterative Krylov methods for large linear systems*, vol. 13 of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2003.
[35] P. Vaněk, M. Brezina, and J. Mandel, *Convergence of algebraic multigrid based on smoothed aggregation.*, UCD/CCD Report 126, University of Colorado at Denver, 1998.
[36] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems.*, Computing, 56 (1996), pp. 179–196.



(a) uniform rectangular grid sequence $\{\mathcal{T}_i^s\}_{i=1}^5$. (b) unstructured grid sequence $\{\mathcal{T}_i^u\}_{i=1}^5$.

FIGURE 5.2. *Asymptotic residual reduction $r_{avg,5}$ for the solution of Poisson's problem. V-cycle V(2,2) with acceleration by the GMRES method.*



(a) Effect of overcorrection for the rectangular grid sequence. SIPG discretization.

(b) Incomplete solve of Poisson's problem using varying numbers of levels. BR2 discretization.

FIGURE 5.3. *Further experiments for the purely elliptic case.*

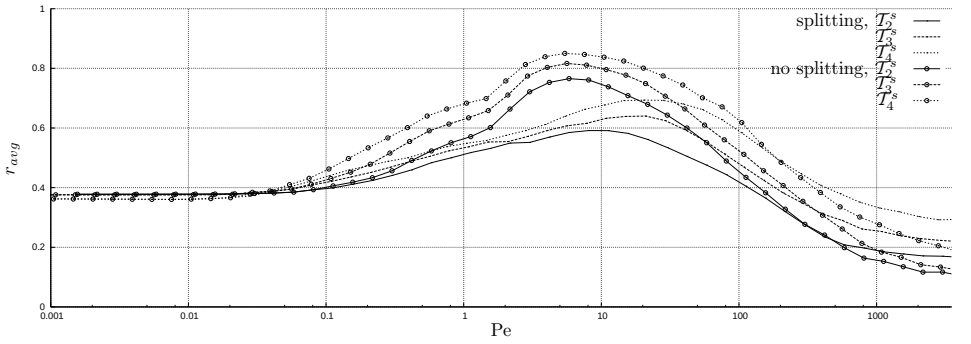


FIGURE 5.4. *Péclet number study for Ritz-Galerkin variational multigrid. BR2 discretization, V(1,1) cycle.Convergence study for varying Péclet number.*

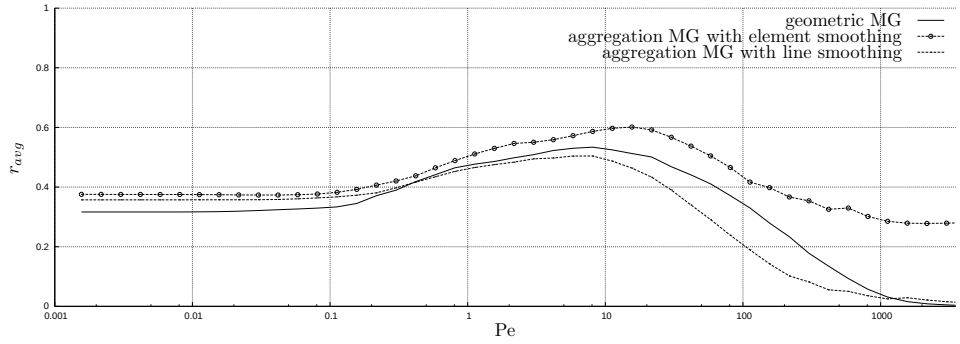
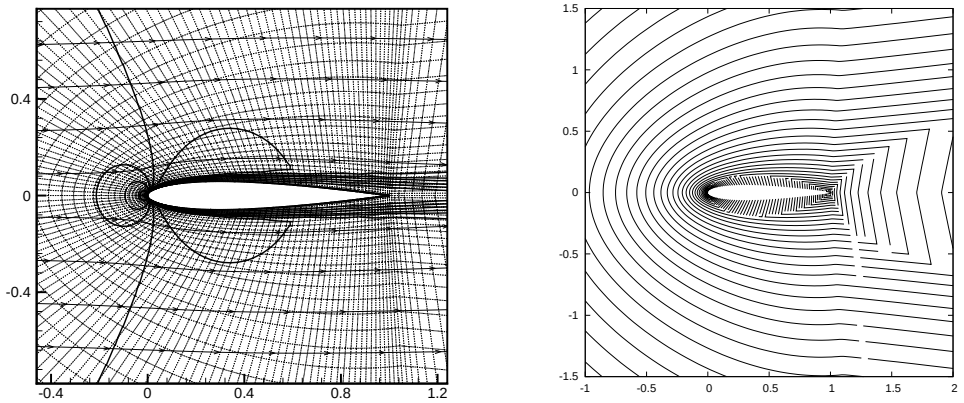
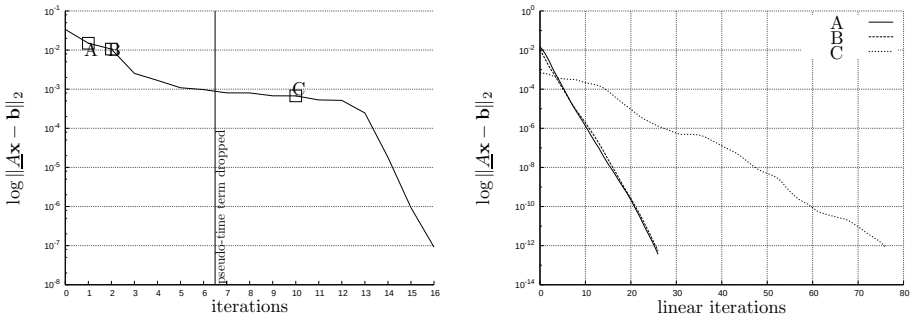


FIGURE 5.5. *Comparison with geometric MG with element-block Gauß-Seidel smoothing and simple injection on mesh $\mathcal{T}_2^s$.SIPG discretization, V(2,2) cycle.*



(a) Solution: Mach number isolines. (b) Result of the line creation heuristic.

FIGURE 5.6. *NACA0012 test case.*



(a) Nonlinear convergence. (b) Linear convergence.

FIGURE 5.7. *Smoothed aggregation multigrid as a preconditioner for the linearized compressible Navier-Stokes equations. Computational mesh with 3072 elements.*